In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institution shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

CONVEX FUNCTIONS

A THESIS

Presented to

The Faculty of the Graduate Division

by

Edward Dixon Lupo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Mathematics

Georgia Institute of Technology

July, 1964

CONVEX FUNCTIONS

Approved:

Date approved by Chairman: January 6, 1965

ACKNOWLEDGMENTS

I wish to express my sincere thanks to Dr. Eric R. Immel, my thesis advisor, for his tireless assistance and expert direction in the preparation of this thesis. In addition I am deeply indebted to Dr. James W. Walker and Dr. Charles H. Braden for their careful criticism of the manuscript and their helpful suggestions to improve it.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The theory of convex functions has assumed substantial importance since its introduction by J.L.W.V. Jensen in 1906 [14].[1] As a part of real-variable theory the theory of convex functions has found uses in the study of inequalities, Orlicz spaces, information theory, and the theory of mathematical statistics and probability. Although frequent use is made of some of the interesting theorems resulting from the theory of convex functions, there seems to be in the literature no unified development of the subject. One of the objectives of this study is to give such a development.

Chapter II is devoted to the development of some of the necessary and sufficient conditions for convexity which are most useful in the application of the theory of convex functions to other fields of mathematics. Utilized in this development is the definition of convexity most commonly found in applications. According to this definition a real-valued function f defined on an interval I in the real line is convex on I if for every x and y in I and every λ in $[0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y) .$$

This, however, is not the definition originally employed by Jensen. The definition introduced in his work, "Sur les fonctions convexes et les

[1]The original development of the theory was due to Jensen, although the convexity inequality (for midpoint convexity) was considered earlier by Hölder [13].

inégalités entre les valeurs moyennes" [14], is more general than the one often considered in applications. A real-valued function f defined on an interval I in the real line is convex on I in the sense of Jensen if for every x and y in I

$$f(\frac{1}{2} x + \frac{1}{2} y) \leq \frac{1}{2} f(x) + \frac{1}{2} f(y) .$$

Hence Chapter II is concerned with a restricted subclass of those real-valued functions defined as convex by Jensen. Herein is a primary source of confusion. Since there are two widely used and accepted definitions of convexity, one might have some difficulty, in connection with a given reference, in determining exactly what properties a given convex function is assumed to possess.

Fortunately, the two definitions are equivalent under relatively mild restrictions. Therefore the confusion, although present, is often of little practical importance in the application of the theory. There is, however, some logical interest in assuming as little as possible, and thus the original definition is often preferred as a starting point. In the logical development of the theory, it is necessary that the distinction between the two definitions and their respective implications, as well as the conditions under which they are equivalent, be made clear. It is with this subject that Chapter III is concerned.

Real-valued functions convex in the sense of Jensen will be called <u>midpoint convex</u> to distinguish them from real-valued functions which satisfy the stronger convexity requirement. Functions of the latter class will be called <u>convex functions</u>. This terminology, though not entirely standard, is nevertheless commonly used in the literature.

Included in Chapter III is the theorem, first proved by Jensen [14], that any continuous midpoint convex function is also convex. Another theorem, also proved by Jensen in the same paper cited above, states that any bounded midpoint convex function on an interval is continuous and thus convex. The remarkable theorem proved by Sierpiński [25] that any measurable midpoint convex function is continuous and hence convex, is also included in Chapter III. These theorems and their respective corollaries establish connections between the two definitions most widely used.

Chapter IV deals with certain important functional equations and properties of functions satisfying these equations. The description elucidates relationships with convexity theory. The first equation discussed is the Cauchy functional equation, $f(x + y) = f(x) + f(y)$. Any real-valued function which satisfies this equation on an interval is midpoint convex, and thus the results of Chapters II and III apply. The proofs of the theorems in Chapter IV, however, are independent, to a large extent, of the preceding chapters. It is shown that any real-valued Lebesgue measurable function f satisfying the Cauchy functional equation is of the form $f(x) = cx$ for some real number c. Also included is an example of a non-Lebesgue-measurable, discontinuous, unbounded real-valued function which satisfies the same functional equation. It is in the description of this rather abstruse example that a Hamel basis for the real numbers is utilized. For completeness, a proof of the existence of a Hamel basis for the real numbers is given in the Appendix, along with some miscellaneous remarks about such a basis.

The second functional equation included in Chapter IV is the equation $f(x + y) = f(x)f(y)$. It is assumed here that f is a complex-valued function with a real interval as its domain. The theorems of this chapter are due to Fréchet [4,5], Hamel [9], Kac [15], and others. The final theorem, based on results of Kac in the paper cited above, asserts that any Lebesgue measurable function satisfying the second functional equation is of the form $f(x) = \exp(ax + ibx)$ for certain real numbers a and b.

CHAPTER II

CONVEXITY

The systematic study of real-valued convex functions was begun by J.L.W.V. Jensen [14] in an important paper published in 1906. In this chapter a more restrictive definition of convex function than that of Jensen is used. The definition used in this chapter implies continuity of the function on the interior of any interval on which the function is convex. Geometrically, a real-valued convex function defined on an interval in the real line is a function for which all chords of the graph lie on or above the graph.

Using this definition of convex function, various criteria for convexity are discussed, along with some examples. Some closure properties for collections of convex functions are noted. Some of the important geometrical properties of convex functions are established. It is shown that a real-valued convex function on an interval in the real line has, at each interior point of the interval, finite one-sided derivatives, and furthermore has, except possibly at a denumerable set of interior points, a finite derivative. Finally, an important integral inequality of Jensen is proved in a form useful in applications.

Definition 2.1. A real-valued function f is convex on a real interval I if I is a subset of the domain of f and for every x and y in I and every λ in $[0,1]$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \ .$$

Example 2.2. Let the function f be defined by $f(x) = x^2$ for all real x. It will be shown that f is convex on the real line, E_1. Let x and y be any two real numbers, and let λ be in $[0,1]$. Thus, $\lambda - \lambda^2 \geq 0$ and $(x - y)^2 = x^2 - 2xy + y^2 \geq 0$. It follows that

$$0 \leq (\lambda - \lambda^2)(x^2 - 2xy + y^2) = (\lambda - \lambda^2)x^2 - 2(\lambda - \lambda^2)xy + (1 - \lambda - 1 + 2\lambda - \lambda^2)y^2$$

and,

$$\lambda^2x^2 + 2\lambda(1 - \lambda)xy + (1 - \lambda)^2y^2 \leq \lambda x^2 + (1 - \lambda)y^2$$

or,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) .$$

Example 2.3. Let f be defined for all $x > 0$ by

$$f(x) = - \log_b x$$

for some base $b > 1$. Then f is convex on $(0, \infty)$.

The preliminary inequality $\lambda(1 - u) \leq 1 - u^\lambda$ for u, λ in $[0,1]$ is used to show that f is convex. Let h be defined by

$$h(u) = u^\lambda$$

for all $u > 0$ and for some fixed $\lambda \in [0, 1]$.

Then

$$h''(u) = (\lambda - 1)\lambda u^{\lambda-2} \leq 0$$

for all $u > 0$. By Taylor's formula there is some $z \in [u, 1]$ such that

$$h(u) = h(1) + h'(1)(u - 1) + \frac{h''(z)}{2!}(u - 1)^2.$$

Thus

$$u^\lambda \le 1 + \lambda(u - 1) ,$$

and

$$\lambda(1 - u) \le 1 - u^\lambda . \qquad (2.3.1)$$

Consider x,y such that $0 < x < y < \infty$. It is necessary to show that

$$\log_b(\lambda x + (1-\lambda)y) \ge \lambda \log_b x + (1-\lambda)\log_b y .$$

By (2.3.1)

$$\lambda(1 - \frac{x}{y}) \le 1 - (\frac{x}{y})^\lambda ,$$

or

$$\lambda(\frac{x}{y}) + (1 - \lambda) \ge (\frac{x}{y})^\lambda ,$$

or

$$\lambda x + (1-\lambda)y \ge x^\lambda y^{1-\lambda} ,$$

and

$$\frac{\lambda x + (1 - \lambda)y}{x^\lambda y^{1-\lambda}} \ge 1 .$$

Thus, since $b > 1$,

$$\log_b\left(\frac{\lambda x + (1 - \lambda)y}{x^\lambda\ y^{1-\lambda}}\right) \geq 0\ ,$$

and,

$$\log_b\ (\lambda x + (1 - \lambda)y) \geq \lambda\ \log_b x + (1 - \lambda)\log_b\ y\ .$$

Note. It follows easily from the above argument that $f(x) = b^x$ is convex for each constant $b > 1$.

Theorem 2.4. Let $f_1, f_2, \ldots, f_n$ be n real-valued functions, each convex on an interval $I \subset E_1$. Let $a_1, a_2, \ldots, a_n$ be n positive real numbers. Then the function f defined by

$$f(x) = \sum_{i=1}^{n} a_i f_i(x)$$

is convex on I.

Proof. The assertion is immediate if $n = 1$. Assume that

$$g_k(x) = \sum_{i=1}^{k} a_i f_i(x)$$

is convex on I for some $k < n$. Then, if $\bar{x} = \lambda x + (1 - \lambda)y$ for $x, y \in I$ and $0 \leq \lambda \leq 1$,

$$g_k(\bar{x}) \leq \lambda g_k(x) + (1 - \lambda)\ g_k(y).$$

Now

$$f_{k+1}(\bar{x}) \leq \lambda f_{k+1}(x) + (1 - \lambda) f_{k+1}(y)\ . \qquad (2.4.1)$$

Thus

$$a_{k+1}f_{k+1}(\bar{x}) \leq \lambda a_{k+1}f_{k+1}(x) + (1-\lambda)a_{k+1}f_{k+1}(y) . \quad (2.4.2)$$

Adding inequalities (2.4.1) and (2.4.2), it follows that

$$g_{k+1}(\bar{x}) \leq \lambda g_{k+1}(x) + (1-\lambda)g_{k+1}(y) .$$

Thus, by induction on n, it follows that

$$f(x) = g_n(x) = \sum_{i=1}^{n} a_i f_i(x)$$

is convex on I. ∎

<u>Theorem 2.5</u>. Let $\{f_t\}_{t\varepsilon T}$ be a collection of real-valued functions each convex on an interval $I \subset E_1$. Let $f(x) = \sup \{f_t(x) : t \varepsilon T\}$ for each $x \varepsilon I$. If f is real-valued on I, then f is convex on I.

<u>Proof</u>. Let x and y be points of I, let $0 \leq \lambda \leq 1$, and define $\bar{x} = \lambda x + (1-\lambda)y$. Then for each $t \varepsilon T$,

$$f_t(\bar{x}) \leq \lambda f_t(x) + (1-\lambda)f_t(y) .$$

Since $f(x)$ is real for each x

$$\lambda f_t(x) \leq \lambda f(x)$$

and,

$$(1-\lambda)\, f_t(y) \leq (1-\lambda)f(y)$$

for each $t \varepsilon T$. Thus

$$\lambda f_t(x) + (1 - \lambda)f_t(y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

for each $t \varepsilon T$. Hence

$$f_t(\bar{x}) \leq \lambda f(x) + (1 - \lambda)f(y)$$

for each $t \varepsilon T$. Therefore

$$f(\bar{x}) = \sup \left\{f_t(\bar{x}) : t \varepsilon T\right\} \leq \lambda f(x) + (1 - \lambda)f(y) ,$$

and f is convex on I.∎

<u>Theorem 2.6</u>. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions each convex on an interval $I \subset E_1$. Let $f(x) = \limsup f_n(x)$ for each x in I. If f is real-valued on I, then f is convex on I.

<u>Proof</u>. Let x and y be points of I, let $0 \leq \lambda \leq 1$, and define

$$\bar{x} = \lambda x + (1 - \lambda)y .$$

Let

$$g_n(x) = \sup \left\{f_n(x), f_{n+1}(x), \ldots.\right\} ,$$

for $n = 1,2,\ldots$. Then g_n is convex on I for $n = 1,2,\ldots$. Thus

$$g_n(\bar{x}) \leq \lambda g_n(x) + (1 - \lambda) g_n(y) .$$

Let $n \rightarrow \infty$, and it follows that

$$\lim_{n\rightarrow\infty} g_n(\bar{x}) \leq \lambda \lim_{n\rightarrow\infty} g_n(x) + (1 - \lambda) \lim_{n\rightarrow\infty} g_n(y)$$

which is the same as

$$\limsup f_n(\bar{x}) \leq \lambda \limsup f_n(x) + (1 - \lambda) \limsup f_n(y).$$

Hence $f = \limsup f_n$ is convex on I.∎

Theorem 2.7. If g is a real-valued function convex on $I \subset E_1$, and f is a real-valued function convex and increasing on an interval containing the range of g, then the composite function $f \circ g$ is convex on I.

Proof. Let x and y be points of I, let $0 \leq \lambda \leq 1$, and define $\bar{x} = \lambda x + (1 - \lambda)y$. Then

$$g(\bar{x}) \leq \lambda g(x) + (1 - \lambda)g(y) .$$

Since f is increasing and convex

$$f \circ g(\bar{x}) = f(g(\bar{x})) \leq f(\lambda g(x) + (1 - \lambda)g(y)) \leq \lambda\ f \circ g(x) + (1-\lambda) f \circ g(y).$$ ∎

Theorem 2.8. Let f be a real-valued function such that f'' exists for each x in an open interval $I \subset E_1$. Then f is convex on I if and only if $f''(x) \geq 0$ for all x in I.

Proof. Assume f is convex on I and suppose $f''(x) < 0$ for some x in I. By definition of f'' at x, for every $h > 0$ it is true that there exists $p > 0$ such that

$$-h < \frac{f'(x + \varepsilon) - f'(x)}{\varepsilon} - f''(x) < h$$

for all $0 < |\varepsilon| \leq p$. Or, for $0 < |\varepsilon| \leq p$,

$$\frac{f'(x + \varepsilon) - f'(x)}{\varepsilon} < f''(x) + h.$$

Thus, if $h = \frac{1}{2} |f''(x)|$, then for $0 < \varepsilon \leq p$

$$\frac{f'(x + \varepsilon) - f'(x)}{\varepsilon} < - h$$

and

$$\frac{f'(x - \varepsilon) - f'(x)}{-\varepsilon} < - h .$$

Thus

$$f'(x + \varepsilon) - f'(x - \varepsilon) \leq -2h\varepsilon$$

for $0 \leq \varepsilon \leq p$. Since f' is continuous on I,

$$\int_0^p f'(x + \varepsilon)d\varepsilon - \int_0^p f'(x - \varepsilon)d\varepsilon \leq - p^2 h ,$$

which, on evaluating the integrals, yields

$$f(x + p) + f(x - p) - 2f(x) \leq - p^2 h < 0 .$$

But, since f is convex on I, taking $\lambda = \frac{1}{2}$,

$$f(x + p) + f(x - p) - 2f(x) \geq 0$$

since x, $x - p$, $x + p$ are in I for the p considered. Thus $f''(x) \geq 0$, since the assumption that $f''(x) < 0$ for some x yields a contradiction.

Now assume $f''(x) \geq 0$. Let x and y be points of I, and let $0 \leq \lambda \leq 1$. Define $\bar{x}$ by

$$\bar{x} = \lambda x + (1 - \lambda)y .$$

By Taylor's formula there exist numbers r_1 and r_2 between x and $\overline{x}$ and y and $\overline{x}$ respectively, such that

$$f(x) = f(\overline{x}) + (x - \overline{x})\, f'(\overline{x}) + \frac{1}{2}(x-\overline{x})^2 f''(r_1) \quad (2.8.1)$$

and

$$f(y) = f(\overline{x}) + (y-\overline{x})\, f'(\overline{x}) + \frac{1}{2}(y-\overline{x})^2 f''(r_2). \quad (2.8.2)$$

Multiplying equation (2.8.1) by λ and equation (2.8.2) by $1 - \lambda$ and adding the results, noting that

$$\lambda(x - \overline{x}) = -(1 - \lambda)(y - \overline{x})$$

and that $f''(r_1)$, $f''(r_2) \geq 0$, it follows that

$$\lambda f(x) + (1 - \lambda)\, f(y) \geq f(\overline{x})$$

or

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1-\lambda) f(y). \blacksquare$$

<u>Remark</u>. If $f''(x) > 0$ on I then the above inequalities are strict, i.e.

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$

unless $x = y$, $\lambda = 1$, or $\lambda = 0$.

<u>Example 2.9</u>. If x, y, a and b are positive then

$$x \log \frac{x}{a} + y \log \frac{y}{b} \geq (x+y) \log \frac{x+y}{a+b}.$$

<u>Proof</u>. By symmetry the assumption that

$$\frac{x}{a} \leq \frac{y}{b}$$

is not restrictive. If $bx = ay$ then the assertion is immediate. Notice that if

$$\frac{x}{a} < \frac{y}{b}$$

then

$$\frac{x}{a} < \frac{x + y}{a + b} < \frac{y}{b} ,$$

since

$$\frac{x + y}{a + b} - \frac{x}{a} = \frac{ax + ay - ax - bx}{a(a + b)} = \frac{ay - bx}{a(a + b)} > 0 ,$$

and

$$\frac{y}{b} - \frac{x + y}{a + b} = \frac{ya + yb - bx - by}{(a + b)b} = \frac{ay - bx}{(a + b)b} > 0 .$$

Noting that

$$\frac{x + y}{a + b} = \frac{\frac{y}{b} - \frac{x + y}{a + b}}{\frac{y}{b} - \frac{x}{a}} \left(\frac{x}{a}\right) + \frac{\frac{x + y}{a + b} - \frac{x}{a}}{\frac{y}{b} - \frac{x}{a}} \left(\frac{y}{b}\right) ,$$

and that $x \log x$ is convex on $(0, \infty)$ since $D_x^2 (x \log x) > 0$ on $(0, \infty)$, it follows that

$$\frac{x + y}{a + b} \log \frac{x + y}{a + b} \leq \frac{x}{a} \left(\frac{\frac{y}{b} - \frac{x + y}{a + b}}{\frac{y}{b} - \frac{x}{a}}\right) \log \frac{x}{a} + \frac{y}{b} \left(\frac{\frac{x+y}{a+b} - \frac{x}{a}}{\frac{y}{b} - \frac{x}{a}}\right) \log \frac{y}{b} .$$

But,

$$\frac{x}{a}\left(\frac{\frac{y}{b}-\frac{x+y}{a+b}}{\frac{y}{b}-\frac{x}{a}}\right)=\frac{x}{a+b},$$

and

$$\frac{y}{b}\left(\frac{\frac{x+y}{a+b}-\frac{x}{a}}{\frac{y}{b}-\frac{x}{a}}\right)=\frac{y}{a+b}.$$

Consequently,

$$(x+y)\log\frac{x+y}{a+b}\leq x\log\frac{x}{a}+y\log\frac{y}{b}. \blacksquare$$

Corollary 2.10. Let f be a real-valued function, positive and twice differentiable on an open interval $I \subset E_1$. The function h, defined by $h(x) = \log f(x)$, is convex on I if and only if $f(x)f''(x)-(f'(x))^2 \geq 0$.

Proof. The function h is convex on I if and only if $h''(x) \geq 0$ on I by Theorem 2.8. The following relations are then equivalent:

$$h''(x) \geq 0;$$

$$D_x(h'(x)) = D_x\left(\frac{1}{f(x)}f'(x)\right) \geq 0;$$

$$\frac{1}{f(x)}f''(x) - \left(f'(x)\frac{1}{f(x)}\right)^2 \geq 0;$$

$$f(x)f''(x) - (f'(x))^2 \geq 0. \blacksquare$$

Theorem 2.11. Let f be a real-valued function defined on an interval $I \subset E_1$. A necessary and sufficient condition that f be convex on I is that, for every $a, b \in I$ $(a < b)$ and for every real r, the function

$$h(x) = f(x) + rx$$

attains its maximum on $[a, b]$ at either a or b.

Proof. Assume f convex on I. Let $[a, b] \subset I$ and let $x \in [a, b]$. Now,

$$x = \frac{b - x}{b - a} a + \frac{x - a}{b - a} b ,$$

and

$$f(x) \leq \frac{b - x}{b - a} f(a) + \frac{x - a}{b - a} f(b) . \qquad (2.11.1)$$

Adding rx to both sides of inequality (2.11.1), it follows that

$$h(x) \leq \frac{b - x}{b - a} h(a) + \frac{x - a}{b - a} h(b) .$$

If $M = \text{Max}\left\{h(a), h(b)\right\}$, then

$$h(x) \leq \frac{b - x}{b - a} M + \frac{x - a}{b - a} M = M$$

for all $x \in [a, b]$. Thus either $h(x) \leq h(a)$ or $h(x) \leq h(b)$ for every x in $[a, b]$.

Now assume that, for every real r, $f(x) + rx$ attains its maximum on $[a, b]$ at either a or b. In the following let r_1 be defined by

$$r_1 = \frac{f(b) - f(a)}{a - b} .$$

Notice that $h(a) = f(a) + r_1 a = f(b) + r_1 b = h(b)$. Thus for every

$x \in [a, b]$

$$h(x) = f(x) + r_1 x \leq f(a) + r_1 a = h(a).$$

Let $\lambda \in [0, 1]$. If $x = \lambda a + (1 - \lambda)b$, then $x \in [a, b]$, and thus,

$$\begin{aligned} f(x) &\leq f(a) + r_1 a - r_1 x = f(a) + r_1 a - r_1 (\lambda a + (1-\lambda)b) \\ &\leq f(a) + (1 - \lambda) r_1 a - (1 - \lambda) r_1 b \\ &\leq f(a) + (1 - \lambda) r_1 (a - b) \\ &\leq f(a) + (1 - \lambda)[f(b) - f(a)] \\ &\leq \lambda f(a) + (1 - \lambda) f(b) . \end{aligned}$$

Similarly, since $f(x) + r_1 x \leq f(b) + r_1 b$, it follows that

$$f(x) \leq \lambda f(b) + (1-\lambda) f(a) ,$$

and thus f is convex on I. ∎

The following lemmas are needed for the proof of an important theorem which provides an additional necessary and sufficient condition for convexity. They are, moreover, important in their own right.

<u>Lemma 2.12</u>. Let f be a real-valued function on an interval $I \subset E_1$. If f is convex on I, and c is in I, the quotient

$$\frac{f(x) - f(c)}{x - c}$$

is an increasing function of x, for x in I.

<u>Proof</u>. Suppose $c < x < y$. By Definition 2.1, the relation

$$x = \frac{y - x}{y - c} c + \frac{x - c}{y - c} y$$

implies that

$$f(x) \leq \frac{y - x}{y - c} f(c) + \frac{x - c}{y - c} f(y)$$

and

$$(y - c)\ f(x) + cf(c) \leq cf(c) + (y - x)\ f(c) + (x - c)\ f(y)$$

or

$$yf(x) - cf(x) + cf(c) - yf(c) \leq cf(c) - xf(c) + xf(y) - cf(y)$$

or

$$(y - c)(f(x) - f(c)) \leq (x - c)(f(y) - f(c)).$$

Thus,

$$\frac{f(x) - f(c)}{x - c} \leq \frac{f(y) - f(c)}{y - c} .$$

Suppose $x < y < c$. By Definition 2.1 the relation

$$y = \frac{c - y}{c - x} x + \frac{y - x}{c - x} c$$

implies that

$$f(y) \leq \frac{c - y}{c - x} f(x) + \frac{y - x}{c - x} f(c) .$$

As before

$$(c - x)\ f(y) + cf(c) \leq cf(c) + (c - y)\ f(x) + (y - x)\ f(c),$$

and

$$yf(x) - yf(c) - cf(x) + cf(c) \leq xf(y) - xf(c) - cf(y) + cf(c) ,$$

or

$$(y - c)(f(x) - f(c)) \leq (x - c)(f(y) - f(c)).$$

Thus, since $(y - c) < 0$ and $(x - c) < 0$,

$$\frac{f(x) - f(c)}{x - c} \leq \frac{f(y) - f(c)}{y - c}.$$

Suppose $x < c < y$. By Definition 2.1, the relation

$$c = \frac{y - c}{y - x} x + \frac{c - x}{y - x} y$$

implies that

$$f(c) \leq \frac{y - c}{y - x} f(x) + \frac{c - x}{y - x} f(y) ,$$

and

$$(y - x) f(c) + cf(c) \leq cf(c) + (y - c) f(x) + (c - x) f(y) .$$

Thus

$$yf(x) - cf(x) - yf(c) + cf(c) \geq xf(y) - xf(c) - cf(y) + cf(c)$$

and

$$(y - c)(f(x) - f(c)) \geq (x - c)(f(y) - f(c)).$$

Finally, since $(y - c) > 0$ and $(x - c) < 0$,

$$\frac{f(x) - f(c)}{x - c} \leq \frac{f(y) - f(c)}{y - c} . \blacksquare$$

Lemma 2.13. Let f be a real-valued function convex on an open interval $I \subset E_1$. Let $y \in I$. If the right-hand derivative $f'_+(y)$ is finite, then for any λ such that $\lambda \leq f'_+(y)$ and all $x \in I$ such that $x > y$, $f(x) \geq f(y) + \lambda(x - y)$. If the left-hand derivative $f'_-(y)$ is finite, then for any λ such that $\lambda \geq f'_-(y)$ and all $x \in I$ such that $x < y$

$$f(x) \geq f(y) + \lambda(x - y) .$$

Proof. An indirect proof is used. Suppose that there exists a point $x \in I$ with $x > y$ and a number $\lambda \leq f'_+(y)$ such that

$$f(x) < f(y) + \lambda(x - y) .$$

Then

$$f'_+(y) \geq \lambda > \frac{f(x) - f(y)}{x - y} .$$

By Lemma 2.12,

$$\frac{f(y + h) - f(y)}{h}$$

is a decreasing function of h as $h \to 0+$. Thus for some positive h,

$$\frac{f(y + h) - f(y)}{h} \geq \lambda .$$

Then,

$$f(y + h) \geq f(y) + h\lambda > f(y) + h\left(\frac{f(x) - f(y)}{x - y}\right), \quad (2.13.1)$$

and

$$f(y) + h\left(\frac{f(x) - f(y)}{x - y}\right) = f(y) - \frac{h}{x-y} f(y) + \frac{h}{x-y} f(x)$$

$$= \frac{x - (y+h)}{x - y} f(y) + \frac{h}{x-y} f(x). \quad (2.13.2)$$

But,

$$y + h = \frac{x - (y + h)}{x - y} y + \frac{h}{x - y} x ;$$

and, by (2.13.1) and (2.13.2),

$$f(y + h) > \frac{x - (y + h)}{x - y} f(y) + \frac{h}{x - y} f(x).$$

This contradicts the fact that f is convex on I.

Then $f(x) \geq f(y) + \lambda(x - y)$ for all $x \in I$ such that $x < y$, if $\lambda \leq f'_+(y)$.

By an exactly analogous argument the second assertion follows. ∎

Lemma 2.14. Let f be a real-valued function convex on an open interval $I \subset E_1$. For each $x \in I$, $f'_+(x)$ and $f'_-(x)$ exist and are finite, and $f'_+(x) \geq f'_-(x)$. This implies, in particular, that f is continuous on I.

Proof. Let $y \in I$ with $y < x$. By Lemma 2.12

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(x + h) - f(x)}{h} \quad (2.14.1)$$

for all $h > 0$ such that $x + h \in I$. If G is defined by

$$G(h) = \frac{f(x + h) - f(x)}{h}$$

for all $h > 0$ such that $x + h \in I$, then by Lemma 2.12, $G(h)$ decreases as $h \longrightarrow 0+$. By (2.14.1), $G(h)$ is bounded below. Thus, $f'_+(x)$ exists and is finite. Similarly $f'_-(x)$ is finite.

It remains to show that $f'_+(x) \geq f'_-(x)$. Suppose that $f'_+(x) < f'_-(x)$. Then for some $h > 0$

$$\frac{f(x - h) - f(x)}{- h} > \frac{f(x + h) - f(x)}{h} .$$

Hence

$$f(x) - f(x - h) > f(x + h) - f(x) ,$$

and

$$f(x) > \frac{f(x + h) + f(x - h)}{2} .$$

This contradicts the fact that f is convex on I. ∎

Theorem 2.15. Let f be a real-valued function on an open interval $I \subset E_1$. Then f is convex on I if and only if for each y in I there exists a real number λ_y such that for all x in I

$$f(x) \geq f(y) + \lambda_y(x - y) .$$

Proof. Assume first that f is convex on I. By Lemma 2.14, $f'_+(y)$

and $f'_-(y)$ are finite for each $y \varepsilon I$. Let $\lambda_y = \frac{1}{2} f'_+(y) + \frac{1}{2} f'_-(y)$. Then $f'_-(y) \leq \lambda_y \leq f'_+(y)$. By Lemma 2.13, $f(x) \geq f(y) + \lambda_y(x - y)$ for all $x \varepsilon I$ such that $x \neq y$. If $x = y$, the assertion of the theorem is trivial.

Now assume that for each $y \varepsilon I$ there exists a number λ_y such that, for all x in I, $f(x) \geq f(y) + \lambda_y(x - y)$. Let x_1 and x_2 be two points of I, and let $y = \lambda x_1 + (1 - \lambda)x_2$ for some $\lambda \varepsilon [0, 1]$. Thus

$$f(x_1) \geq f(y) + \lambda_y(x_1 - y) ,$$

and

$$f(x_2) \geq f(y) + \lambda_y(x_2 - y) .$$

Hence,

$$f(y) \leq f(x_1) - \lambda_y(x_1 - y) ,$$

and

$$f(y) \leq f(x_2) - \lambda_y(x_2 - y) .$$

Now,

$$x_1 - y = (1 - \lambda)x_1 - (1 - \lambda)x_2 = (1 - \lambda)(x_1 - x_2) ,$$

and

$$x_2 - y = -\lambda x_1 + \lambda x_2 = - \lambda(x_1 - x_2) .$$

Thus,

$$f(y) \leq f(x_1) - \lambda_y(1 - \lambda)(x_1 - x_2) , \qquad (2.15.1)$$

and

$$f(y) \leq f(x_2) + \lambda_y \lambda \ (x_1 - x_2) \ . \qquad (2.15.2)$$

Multiply (2.15.1) by λ and (2.15.2) by $1 - \lambda$ and add. Thus

$$f(y) \leq \lambda f(x_1) + (1 - \lambda) \ f(x_2) \ . \blacksquare$$

Lemma 2.16. Let f be real-valued and convex in an interval $I \subset E_1$. Let a and b be interior points of I, with $a < b$. Then

$$f'_+(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'_-(b) \ .$$

Proof. By Lemma 2.12, for each $\varepsilon > 0$ such that $a < a + \varepsilon \leq b$

$$\frac{f(a + \varepsilon) - f(a)}{\varepsilon} \leq \frac{f(b) - f(a)}{b - a} \, .$$

Thus, since $f'_+(a)$ is finite, it follows that

$$f'_+(a) \leq \frac{f(b) - f(a)}{b - a} \, .$$

Again by Lemma 2.12, for each $\varepsilon < 0$ such that $a \leq b + \varepsilon < b$,

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(b + \varepsilon)}{- \varepsilon} = \frac{f(b + \varepsilon) - f(b)}{\varepsilon} \, .$$

Thus, since $f'_-(b)$ is finite, it follows that

$$\frac{f(b) - f(a)}{b - a} \leq f'_-(b) \ . \blacksquare$$

Theorem 2.17. If f is a real-valued and convex in an interval $I \subset E_1$, then the set of points of I at which f is non-differentiable is countable.

Proof. Define the set A by

$$A = \{x : x \text{ is interior to } I \text{ and } f'(x) \text{ does not exist}\}.$$

For each $x \varepsilon I$, $f'_+(x) \geq f'_-(x)$ by Lemma 2.14. If $x \varepsilon A$, then $f'_+(x) > f'_-(x)$. For each $x \varepsilon A$ let I_x be the open interval $(f'_-(x),\ f'_+(x))$. Suppose that x_1 and x_2 are points of A with $x_1 < x_2$. By Lemma 2.16,

$$f'_+(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'_-(x_2) .$$

Thus

$$f'_-(x_1) < f'_+(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'_-(x_2) < f'_+(x_2) .$$

It follows from these inequalities that $I_{x_1} \cap I_{x_2} = \emptyset$. Notice that, for each $x \varepsilon A$, the interval I_x is nonempty. Thus the intervals of the collection $\{I_x\}$ are open, nonempty, and pairwise disjoint. It is easily shown that the collection is countable since in each set I_x of the collection there is a rational number and the rationals are countable. It follows that the set A itself is countable. ∎

Note. This theorem is due to L. Galvani [6].

Theorem 2.18. (Jensen's integral inequality) Let f be a real-valued function convex on E_1. Let g be a Lebesgue measurable function finite real-valued almost everywhere on an interval $[a, b] \subset E_1$. Finally, let w be a non-negative function such that the functions w and wg are Lebesgue summable on $[a, b]$ and

$$\int_a^b w(x)\, dx > 0 .$$

Then

$$f\left(\frac{\int_a^b g(x)\, w(x)\, dx}{\int_a^b w(x)\, dx}\right) \leq \frac{\int_a^b f(g(x))\, w(x)\, dx}{\int_a^b w(x)\, dx} . \qquad (2.18.1)$$

Note. If the mean value of $g(x)$ on $[a, b]$ relative to the weight function $w(x)$ is defined by

$$M[g] = \frac{\int_a^b g(x)\, w(x)\, dx}{\int_a^b w(x)\, dx} ,$$

then the Jensen inequality in the form (2.18.1) asserts that

$$f(M[g]) \leq M[f \circ g] ,$$

where $f \circ g$ denotes the composition of f and g.

Proof. Since f is convex on E_1, for every $y \in E_1$ there is, noting Theorem 2.15, a real number $\lambda(y)$ such that

$$\lambda(y)(z - y) \leq f(z) - f(y) \qquad (2.18.2)$$

for all $z \in E_1$. Let $\tilde{g}$ be defined by

$$\begin{aligned} \tilde{g}(x) &= g(x) \quad \text{if } |g(x)| < \infty \\ &= 0 \quad \text{otherwise} , \end{aligned}$$

and let

$$y_1 = \frac{\int_a^b g(x)\, w(x)\, dx}{\int_a^b w(x)\, dx} .$$

By hypothesis, y_1 is finite. The functions $\tilde{g}$ and $f \circ \tilde{g}$ are finite-valued and $f \circ \tilde{g} = f \circ g$ a.e. on $[a, b]$. From (2.18.2)

$$\lambda(y_1)(\tilde{g}(x) - y_1) \leq (f \circ \tilde{g})(x) - f(y_1) \qquad (2.18.3)$$

for each x in $[a, b]$. Let $\tilde{w}$ be defined

$$\begin{aligned} \tilde{w}(x) &= w(x) \quad \text{if} \quad |w(x)| < \infty \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

From (2.18.3)

$$(f \circ \tilde{g})(x) \geq \lambda(y_1)(\tilde{g}(x) - y_1) + f(y_1).$$

Thus

$$0 < (f \circ \tilde{g})^-(x) \leq - \lambda(y_1)(\tilde{g}(x) - y_1) - f(y_1)$$

for each $x \varepsilon [a, b]$ such that $(f \circ \tilde{g})(x) < 0$, since

$$\begin{aligned} (f \circ \tilde{g})^-(x) &= - (f \circ \tilde{g})(x) \quad \text{if} \quad (f \circ \tilde{g})(x) < 0 , \\ &= 0 \qquad\qquad \text{otherwise.} \end{aligned}$$

Hence

$$0 \leq (f \circ \tilde{g})^-(x) \leq - \lambda(y_1)(\tilde{g}(x)-y_1) - f(y_1),$$

for each $x \varepsilon [a, b]$. Since $\tilde{w}(x) \geq 0$,

$$0 \leq (f \circ \tilde{g})^{-}(x)\tilde{w}(x) \leq -\lambda(y_1)\,\tilde{w}(x)(\tilde{g}(x) - y_1) - \tilde{w}(x)\,f(y_1)\,.$$

Thus, since $\tilde{g}\tilde{w}$ and $\tilde{w}$ are summable, $(f \circ \tilde{g})^{-}\tilde{w}$ is summable, and it follows that

$$\int_a^b (f \circ g)(x)\,w(x)\,dx = \int_a^b (f \circ \tilde{g})(x)\,\tilde{w}(x)\,dx$$

exists, possibly as $+\infty$.

Again from (2.18.3) and the fact that $\tilde{w}(x) \geq 0$ it follows that

$$\lambda(y_1)\,\tilde{w}(x)(\tilde{g}(x) - y_1) \leq \tilde{w}(x)(f \circ \tilde{g})(x) - \tilde{w}(x)\,f(y_1) \quad (2.18.4)$$

for each $x \varepsilon [a, b]$. Since for both sides of inequality (2.18.4) the Lebesgue integral exists

$$\lambda(y_1)\left(\int_a^b \tilde{w}(x)\,\tilde{g}(x)\,dx - y_1\int_a^b \tilde{w}(x)\,dx\right) \leq$$

$$\int_a^b (f \circ \tilde{g})(x)\,\tilde{w}(x)\,dx - f(y_1)\int_a^b \tilde{w}(x)\,dx.$$

However,

$$y_1\int_a^b \tilde{w}(x)\,dx = \int_a^b w(x)\,g(x)\,dx = \int_a^b \tilde{w}(x)\,\tilde{g}(x)\,dx\,.$$

Hence,

$$\int_a^b (f \circ g)(x)\,w(x)\,dx - f(y_1)\int_a^b w(x)\,dx \geq$$

$$\lambda(y_1)\left(\int_a^b w(x)\,g(x)\,dx - y_1\int_a^b w(x)\,dx\right) = 0.$$

Since

$$\int_a^b w(x)\,dx > 0 ,$$

it follows that

$$f(y_1) \le \frac{\int_a^b (f \circ g)(x)\, w(x)\, dx}{\int_a^b w(x)\, dx} ,$$

which is the assertion of the theorem. ∎

Remark. Jensen first proved this theorem under somewhat more restrictive hypotheses in 1906 (cf. Jensen [14]). An elaborate proof is given in Natanson [23 (Vol. II, pp. 46 and 47)] with the restriction that $[a, b]$ be a bounded interval.

CHAPTER III

MIDPOINT CONVEXITY AND RELATIONS BETWEEN LEBESGUE MEASURABILITY AND CONVEXITY

In this chapter the concept of a real-valued midpoint convex function is introduced. This is the concept originally used by Jensen, as was pointed out briefly in the introduction to Chapter II. The term midpoint convexity is used to distinguish this less restrictive property from the property of convexity used in Chapter II. As was proved in Chapter II, any real-valued convex function on an interval in the real line is necessarily continuous at all interior points of the interval. An example discussed in Chapter IV shows that this is not the case for midpoint convex functions. However in the present chapter it is shown that a continuous midpoint convex function on an open interval is convex in the sense of Chapter II. This result was first proved by Jensen [14]. It is proved that, under surprisingly mild conditions (such as Lebesgue measurability), a real-valued function which is midpoint convex on an interval is continuous and hence, in fact, convex on the interval.

Definition 3.1. A real-valued function f is midpoint convex on the interval $I \subset E_1$, if I is a subset of the domain of f and for every x and y in I,

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} .$$

Theorem 3.2. Let the interval $I \subset E_1$ be a subset of the domain of the real-valued function f. Then f is midpoint convex on I if and only if

$$f\left(\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}\right) \leq \frac{\sum_{i=1}^{n} p_i f(x_i)}{\sum_{i=1}^{n} p_i} \qquad (3.2.1)$$

for any set $\{x_i\}$ of n elements of I and any set $\{p_i\}$ of n nonnegative rational numbers not all zero.

Remark. The inequality concerned (3.2.1) is sometimes referred to as Jensen's summation inequality (cf. Natanson [23], Vol. II, p. 44).

Proof. The fact that condition (3.2.1) implies that f is midpoint convex is an immediate consequence of Definition 3.1 as seen by letting $n = 2$ and $p_1 = p_2 = 1$. The proof that midpoint convexity of f on I implies condition (3.2.1) is given in three parts. Assume in the following that f is midpoint convex on the interval I.

Part (i). If $n = 2^m$ for some integer $m > 0$, then

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \leq \frac{1}{n}\sum_{i=1}^{n} f(x_i) \ . \qquad (3.2.2)$$

The proof of the assertion is by induction on m. For $m = 1$ ($n = 2$) the assertion reduces to the hypothesis of midpoint convexity.

It remains to show that the proposition is true for $m = k + 1$ if it is true for $m = k$. Consider any set $\{x_1, \ldots, x_{2^k}, x_{2^k+1}, \ldots, x_{2^{k+1}}\}$ of 2^{k+1} points in I. Let

$$\bar{x} = \frac{1}{2^k} \sum_{i=1}^{2^k} x_i \ , \quad \tilde{x} = \frac{1}{2^k} \sum_{i=1}^{2^k} x_{2^k+i} \ .$$

Both $\bar{x}$ and $\tilde{x}$ belong to I; and, since f is midpoint convex on I,

$$f\left(\frac{\bar{x}+\tilde{x}}{2}\right) \le \frac{f(\bar{x}) + f(\tilde{x})}{2} \ . \tag{3.2.3}$$

Now

$$\frac{\bar{x}+\tilde{x}}{2} = \frac{1}{2^{k+1}} \sum_{i=1}^{2^{k+1}} x_i \ .$$

Hence, from (3.2.3),

$$f\left(\frac{1}{2^{k+1}} \sum_{i=1}^{2^{k+1}} x_i\right) = f\left(\frac{\bar{x}+\tilde{x}}{2}\right) \le \frac{1}{2}\left(f(\bar{x}) + f(\tilde{x})\right) . \tag{3.2.4}$$

But, if the assertion is true for $m = k$, then

$$f(\bar{x}) = f\left(\frac{1}{2^k} \sum_{i=1}^{2^k} x_i\right) \le \frac{1}{2^k} \sum_{i=1}^{2^k} f(x_i) \ ,$$

and similarly

$$f(\tilde{x}) \le \frac{1}{2^k} \sum_{i=1}^{2^k} f(x_{2^k+i}) \ .$$

These inequalities, together with (3.2.4), imply that the assertion must also then be true for $m = k + 1$.

Thus, the assertion of Part (i) is proved for $n = 2^m$ and m a positive integer.

Part (ii). If n is any positive integer then inequality (3.2.2) holds.

Let $\{x_1, x_2, \ldots, x_n\}$ be a set of n points of I, and let m be such that $2^m > n$. Define $\bar{\bar{x}}$ by

$$\bar{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} x_i \; .$$

Note that $\bar{\bar{x}} \in I$. Now consider the set of 2^m points of I,

$$\left\{x_1', x_2', \ldots, x_n', x_{n+1}', \ldots, x_{2^m}'\right\},$$

where $x_i' = x_i$ for $1 \le i \le n$ and $x_i' = \bar{\bar{x}}$ for $n < i \le 2^m$. It follows that

$$\bar{\bar{x}} = \frac{n\bar{\bar{x}} + (2^m - n)\bar{\bar{x}}}{2^m} = \frac{\sum_{i=1}^{n} x_i + \sum_{i=n+1}^{2^n} \bar{\bar{x}}}{2^m} = \frac{1}{2^m} \sum_{i=1}^{2^m} x_i' \; .$$

By Part (i)

$$f(\bar{\bar{x}}) \le \frac{1}{2^m} \sum_{i=1}^{2^m} f(x_i') = \frac{\sum_{i=1}^{n} f(x_i) + (2^m - n)\; f(\bar{\bar{x}})}{2^m}$$

or,

$$f(\bar{\bar{x}}) - \frac{(2^m - n)\; f(\bar{\bar{x}})}{2^m} \le \frac{1}{2^m} \sum_{i=1}^{n} f(x_i) \; .$$

By multiplying by $\frac{1}{n}(2^m)$, it follows that

$$f(\bar{\bar{x}}) \leq \frac{1}{n} \sum_{i=1}^{n} f(x_i) \ ,$$

which was to be proved.

Part (iii). Let $p_1, p_2, \ldots, p_n$ be nonnegative rational numbers, not all zero. Then it must be proved that

$$f\left(\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}\right) \leq \frac{\sum_{i=1}^{n} p_i f(x_i)}{\sum_{i=1}^{n} p_i} \ .$$

It suffices to consider the case that $p_i > 0$ for $i = 1,2,\ldots,n$. For each $i = 1,2,\ldots, n$, let $p_i = r_i/s_i$ where r_i and s_i are positive integers and each r_i/s_i is in lowest terms. Let

$$p = \frac{1}{\text{l.c.m.} \ \{s_1, s_2, \ldots, s_n\}} \ ,$$

where $\text{l.c.m.}\{s_1, s_2, \ldots, s_n\}$ is the lowest common multiple of the numbers $s_1, s_2, \ldots, s_n$. It follows that there exists positive integers $n_1, n_2, \ldots, n_n$ for which $p_i = n_i p$ for $i = 1,2,\ldots, n$.

Define $\bar{n}$ by

$$\bar{n} = \sum_{i=1}^{n} n_i \ .$$

Then,

$$\sum_{i=1}^{n} p_i x_i = p \sum_{i=1}^{n} n_i x_i = p \sum_{j=1}^{\bar{n}} y_j \ ,$$

where $y_j = x_1$, for $j = 1,2,\dots, n_1$; $y_j = x_2$ for $j = n_1+1,\dots, n_1+n_2;\dots;$ $y_j = x_n$ for $j = \bar{n} - n_n + 1,\dots, \bar{n}$. Since

$$\sum_{i=1}^{n} p_i = p\,\bar{n}\,,$$

it follows that

$$\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i} = \frac{1}{\bar{n}} \sum_{i=1}^{\bar{n}} y_j\,.$$

By Part (ii),

$$f\left(\frac{1}{\bar{n}} \sum_{j=1}^{\bar{n}} y_j\right) \le \frac{1}{\bar{n}} \sum_{j=1}^{\bar{n}} f(y_j) =$$

$$\frac{p}{p}\,\frac{\sum_{i=1}^{n} n_i f(x_i)}{\bar{n}} = \frac{\sum_{i=1}^{n} p_i f(x_i)}{\sum_{i=1}^{n} p_i}\,. \blacksquare$$

<u>Note</u>. The number $\dfrac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}$ is a point of I.

It has already been shown that, if f is convex on I, then f is continuous on the interior of I. Thus, if f is convex on an open interval I, then f is continuous and midpoint convex. A converse to this last statement is proved next.

Theorem 3.3. Let f be a real-valued continuous function on an interval $I \subset E_1$. If f is midpoint convex on I, then, whenever $x_1, x_2, \ldots, x_n$ are points of I and $p_1, p_2, \ldots, p_n$ are nonnegative real numbers, not all zero, it is true that

$$f\left(\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}\right) \leq \frac{\sum_{i=1}^{n} p_i f(x_i)}{\sum_{i=1}^{n} p_i} . \tag{3.3.1}$$

In particular, f is then convex on I.

Proof. It suffices to suppose $p_i > 0$ for $i = 1,2,\ldots, n$. For each p_i and each integer $j \geq 1$ consider the interval $(p_i, p_i + \frac{1}{j})$. Select one rational number, p_{ij}, contained in $(p_i, p_i + \frac{1}{j})$. Then by Theorem 3.2, for each $j = 1,\ldots,$

$$f\left(\frac{\sum_{i=1}^{n} p_{ij} x_i}{\sum_{i=1}^{n} p_{ij}}\right) \leq \frac{\sum_{i=1}^{n} p_{ij} f(x_i)}{\sum_{i=1}^{n} p_{ij}} . \tag{3.3.2}$$

Since $p_i < p_{ij} < p_i + \frac{1}{j}$, it follows that

$$\lim_{j \to \infty} \frac{\sum_{i=1}^{n} p_{ij} x_i}{\sum_{i=1}^{n} p_{ij}} = \frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i} , \tag{3.3.3}$$

and

$$\lim_{j\to\infty} \frac{\sum_{i=1}^{n} p_{ij} f(x_i)}{\sum_{i=1}^{n} p_{ij}} = \frac{\sum_{i=1}^{n} p_i f(x_i)}{\sum_{i=1}^{n} p_i} . \qquad (3.3.4)$$

Since f is continuous on I, and the point $\sum_{i=1}^{n} p_i x_i \Big/ \sum_{i=1}^{n} p_i$ belongs to I,

$$\lim_{j\to\infty} f\left(\frac{\sum_{i=1}^{n} p_{ij} x_i}{\sum_{i=1}^{n} p_{ij}}\right) = f\left(\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}\right). \qquad (3.3.5)$$

Thus, considering the limits (3.3.3), (3.3.4), and (3.3.5) and the inequality (3.3.2), it follows that

$$f\left(\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}\right) \leq \frac{\sum_{i=1}^{n} p_i f(x_i)}{\sum_{i=1}^{n} p_i} . \blacksquare$$

Theorem 3.4. If the real-valued function f is midpoint convex on a finite closed interval $[a, b]$ and if f is bounded above on (a, b), then f is continuous on (a, b).

Proof. Let M be an upper bound of f on (a, b). Let $\varepsilon > 0$ be given, and consider a fixed point x in (a, b). Let n be a positive integer such that

$$\frac{M - f(x)}{n} < \varepsilon ,$$

and fix such an n in what follows. Let $\delta_n > 0$ be such that $a < x \pm n\delta_n < b$ and consider all δ such that $|\delta| < \delta_n$. Then $a < x + n\delta < b$, and

$$f(x + \delta) = f(\tfrac{1}{n}(x + n\delta) + (\tfrac{n-1}{n})x) \leq \tfrac{1}{n} f(x + n\delta) + \tfrac{n-1}{n} f(x),$$

by Theorem 3.2. Thus

$$f(x + \delta) - f(x) \leq \tfrac{1}{n} (f(x + n\delta) - f(x)) . \qquad (3.4.1)$$

A similar argument shows that

$$f(x - \delta) - f(x) \leq \tfrac{1}{n} (f(x - n\delta) - f(x)). \qquad (3.4.2)$$

Since f is midpoint convex,

$$2f(x) \leq f(x + \delta) + f(x - \delta) ,$$

and thus

$$f(x + \delta) - f(x) \geq f(x) - f(x - \delta). \qquad (3.4.3)$$

Combining (3.4.1), (3.4.2), and (3.4.3),

$$\frac{f(x + n\delta) - f(x)}{n} \geq f(x + \delta) - f(x) \geq f(x) - f(x - \delta)$$
$$\geq \frac{f(x) - f(x - n\delta)}{n} .$$

Since $f(x + n\delta) \leq M$ and $f(x - n\delta) \leq M$,

$$\frac{M - f(x)}{n} \geq f(x + \delta) - f(x) \geq f(x) - f(x - \delta) \geq \frac{f(x)-M}{n}$$

and

$$|f(x + \delta) - f(x)| \leq \frac{M - f(x)}{n} < \varepsilon, \quad |\delta| < \delta_n .$$

Thus f is continuous at each $x \in (a, b)$. ∎

Remark. The above theorem is due to Jensen [14].

Corollary 3.5. Let the real-valued function f be midpoint convex on a finite closed interval $[a, b]$ and bounded above in some interval $(c, d) \subset [a, b]$. Then f is bounded above on $[a, b]$ and hence continuous on (a, b).

Proof. Let M be an upper bound of f on (c, d). If $a = c$ and $d = b$, the assertion is evident. If $a < c$, let $x \in (a, c)$. Define the function g by

$$g(\lambda) = a + \lambda(x - a).$$

Notice that g is continuous and there is some λ_o such that $c < g(\lambda_o) < d$. Thus there is a neighborhood of λ_o, $N(\lambda_o)$, such that $c < g(\lambda) < d$ for all $\lambda \in N(\lambda_o)$. Hence there exists a positive rational number r such that $c < a + r(x - a) < d$. Thus there are positive integers n and m such that

$$c < a + \frac{n}{m}(x - a) < d .$$

Since $a < c$ it follows that $m < n$.

Let

$$y = a + \frac{n}{m}(x - a) .$$

Then

$$x = \frac{m}{n} y + \frac{n - m}{n} a .$$

Since f is midpoint convex on $[a, b]$,

$$f(x) \leq \frac{m}{n} f(y) + \frac{n - m}{n} f(a) .$$

Noting that $c < y < d$ and $f(y) \leq M$, it follows that

$$f(x) \leq \frac{m}{n} M + \frac{n - m}{n} f(a) \leq \operatorname{Max}\left\{M, f(a)\right\} .$$

Thus f is bounded above on (a, c) and thus also on $[a, c]$.

If $d < b$, a similar argument shows that f is bounded above on $[d, b]$. Hence f is bounded above on $[a, b]$. ∎

The following lemma is used to prove the surprising result of Sierpiński [25] mentioned in Chapter I.

<u>Lemma 3.6</u>. Let the function g be defined by $g(x) = ax + b$ for all real x, where a and b are real numbers, and let A be a Lebesgue measurable set in E_1. Then $g(A) = \{g(x) : x \in A\}$ is Lebesgue measurable and $\mu g(A) = |a| \mu A$, where μ is Lebesgue measure in E_1.

<u>Proof</u>. If $a = 0$, or if A is empty the assertion is obvious. Let $a \neq 0$ and A be nonempty. Let m be one-dimensional Lebesgue outer measure. It is proved first (in Part (i)) that $mg(A) = |a| mA$, and then (in Part (ii)) that $g(A)$ is measurable.

Part (i). Let A be a bounded open interval (c, d). Then $mA = d - c$. If $a > 0$, $x \in A$ if and only if $ac + b < g(x) < ad + b$. If $a < 0$, $x \in A$

if and only if $ad + b < g(x) < ac + b$. Thus in either case

$$\mu g(A) = |a|(d - c) . \tag{3.6.1}$$

Let A be any subset of E_1. For each $\varepsilon > 0$ there exists an open set G_ε such that $A \subset G_\varepsilon$ and

$$\mu G_\varepsilon \leq mA + \varepsilon , \tag{3.6.2}$$

since mA is the greatest lower bound of the measures of open sets containing A. Since G_ε is open there exists a countable collection $\{J_n\}_1^\infty$ of pairwise disjoint open intervals such that

$$G_\varepsilon = \bigcup_{n=1}^{\infty} J_n .$$

By (3.6.1)

$$\mu g(J_n) = |a| \mu J_n . \quad \text{Since} \quad A \subset G_\varepsilon ,$$

$$g(A) \subset g(G_\varepsilon) = \bigcup_{n=1}^{\infty} g(J_n) .$$

Thus

$$m g(A) \leq m g(G_\varepsilon) = \mu g(G_\varepsilon) = \mu \left(\bigcup_{n=1}^{\infty} g(J_n) \right)$$

$$= \sum_{n=1}^{\infty} \mu g(J_n) = \sum_{n=1}^{\infty} |a| \mu J_n .$$

It follows that

$$m g(A) \leq \sum_{n=1}^{\infty} |a| \mu J_n = |a| \mu \bigcup_{n=1}^{\infty} J_n = |a| \mu G_\varepsilon \leq |a| (mA + \varepsilon),$$

and thus

$$m\,g(A) \leq |a|\, m A . \tag{3.6.3}$$

Now

$$g^{-1}(y) = \frac{1}{a} y - \frac{b}{a}$$

and

$$g^{-1}(g(A)) = A .$$

By (3.6.3), and applying the result to g^{-1},

$$m A = m g^{-1}(g(A)) \leq \left|\frac{1}{a}\right| m\, g(A) ,$$

and

$$m\,g(A) \geq |a|\, m A .$$

This inequality and (3.6.3) imply that

$$m\,g(A) = |a|\, m A .$$

Part (ii). To show that $g(A)$ is measurable if A is measurable it will be proved that, for any set T of real numbers,

$$m T = m(T \cap g(A)) + m(T \cap g(A)^c). \tag{3.6.4}$$

The following identities will be used.

$$g(T \cap A) = g(T) \cap g(A) , \tag{3.6.5}$$

$$g(T \cap A^c) = g(T) \cap g(A^c) , \tag{3.6.6}$$

and

$$g(A^c) = g(A)^c . \tag{3.6.7}$$

The identity (3.6.7) is proved below. The proofs of (3.6.5) and (3.6.6) are omitted.

For $y \in g(A^c)$ there exists an $\bar{x} \in A^c$ such that $y = a\bar{x} + b$. Suppose also that $y \in g(A)$. Then there is an $x \in A$ such that $y = ax + b$. Thus $\bar{x} = x \in A$, a contradiction.
Thus

$$g(A^c) \subset g(A)^c .$$

If $y \in g(A)^c$, let

$$x = \frac{1}{a} y - \frac{b}{a} .$$

If $x \in A$, then $y \in g(A)$ since $y = g(x)$. Thus $x \in A^c$ and $y \in g(A^c)$.
Hence

$$g(A)^c \subset g(A^c) ,$$

and (3.6.7) follows.

Since A is measurable, for any set $T \subset E_1$

$$mT = m(T \cap A) + m(T \cap A^c) .$$

Now

$$mg(T \cap A) = |a| \, m(T \cap A) ,$$

and

$$mg(T \cap A^c) = |a| \, m(T \cap A^c) ,$$

by the results of Part (i). Thus

$$mT = \frac{1}{|a|}\,(m\,g\,(T \cap A) + mg(T \cap A^c)).$$

By the identities (3.6.5), (3.6.6), and (3.6.7),

$$mT = \frac{1}{|a|}\,m(g(T) \cap g(A)) + m(g(T) \cap g(A)^c). \qquad (3.6.8)$$

Since (3.6.8) is true for any set $T \subset E_1$, it must in particular be true for $g^{-1}(T)$, and thus

$$\begin{aligned} mg^{-1}(T) &= \frac{1}{|a|}\left(m\left(g\ (g^{-1}(T)) \cap g(A)\right) + m\left(g\ (g^{-1}(T)) \cap g(A)^c\right)\right) \\ &= \frac{1}{|a|}\ m\left((T \cap g(A)) + m(T \cap g(A)^c)\right). \end{aligned}$$

But,

$$m\,g^{-1}\,(T) = \frac{1}{|a|}\,m\,T$$

by the results of Part (i). Thus

$$mT = m(T \cap g(A)) + m(T \cap g(A)^c).$$

This establishes (3.6.4).∎

<u>Theorem 3.7</u>. Let f be a real-valued function midpoint convex on an interval (a, b), possibly unbounded. Then f is continuous on (a, b) if and only if f is Lebesgue measurable on (a, b).

<u>Proof</u>. The fact that any continuous function is Lebesgue measurable is well known and is not proved here. What is proved is that a Lebesgue measurable midpoint convex function f is continuous. The proof is indirect.

Suppose that there is a point x_o in (a, b) such that f is not continuous at x_o. Let $h > 0$ be such that $(x_o - 2h, x_o + 2h) \subset (a,b)$. Then by Corollary 3.5, f is unbounded in $I = (x_o - h, x_o + h)$. Thus for each positive integer n there is a point $x_n \varepsilon I$ such that $f(x_n) > n$. If x' and x'' are any two points in $(x_o - 2h, x_o + 2h)$ such that

$$\frac{1}{2}(x' + x'') = x_n , \tag{3.7.1}$$

then at least one of the inequalities

$$f(x') > n, \quad f(x'') > n , \tag{3.7.2}$$

must hold. For, if $f(x') \leq n$, and $f(x'') \leq n$ then by midpoint convexity

$$f(x_n) = f(\tfrac{1}{2} x' + \tfrac{1}{2} x'') \leq \tfrac{1}{2} f(x') + \tfrac{1}{2} f(x'') \leq n,$$

which would contradict the inequality $f(x_n) > n$.

Let n be some positive integer and define M_n by

$$M_n = \left\{x \varepsilon (x_n - h, x_n + h) : f(x) > n\right\}.$$

Then since f is a measurable function on

$$(x_n - h, x_n + h) \subset (x_o - 2h, x_o + 2h) \subset (a, b)$$

M_n is a measurable set of real numbers. Let M_n^+ be the subset of M_n to the right of x_n; that is, $M_n^+ = M_n \cap (x_n, b)$. Let M_n^- be the subset of M_n to the left of x_n; that is, $M_n^- = M_n \cap (a, x_n)$. Then

$$\mu M_n = \mu M_n^+ + \mu M_n^- . \tag{3.7.3}$$

Let J_n^+ and J_n^- be subintervals of the interval $(x_n - h,\ x_n + h)$ defined by $J_n^+ = (x_n,\ x_n + h)$ and $J_n^- = (x_n - h,\ x_n)$. By construction $M_n^+ \subset J_n^+$, $M_n^- \subset J_n^-$ and

$$(M_n^+ \cup M_n^-) \subset (J_n^+ \cup J_n^-) \subset (x_o - 2h,\ x_o + 2h) \subset (a,b). \quad (3.7.4)$$

Notice also that

$$J_n^+ - M_n^+ = \left\{ x \varepsilon J_n^+ :\ f(x) \le n \right\}.$$

It will be demonstrated that

$$\mu(J_n^+ - M_n^+) \le \mu M_n^- .$$

Define the function g_n by

$$g_n(x) = -x + 2x_n .$$

Now $x \varepsilon J_n^+ - M_n^+$ implies that $f(x) \le n$. But for such x, since $\frac{1}{2}x + \frac{1}{2}g_n(x) = x_n$, it follows from (3.7.2) that $f(g_n(x)) > n$ for each $x \varepsilon J_n^+ - M_n^+$. Thus $g_n(x) \ \varepsilon\ M_n^-$ for each $x \varepsilon J_n^+ - M_n^+$ since $g_n(x) < x_n$ for $x > x_n$. Hence $g_n(J_n^+ - M_n^+) \subset M_n^-$. By Lemma 3.6, $\mu g_n(J_n^+ - M_n^+) = \mu(J_n^+ - M_n^+)$ since the absolute value of the slope of g_n is unity. Therefore

$$\mu(J_n^+ - M_n^+) \le \mu M_n^- .$$

Using equation (3.7.3) and the preceding inequality,

$$\mu M_n \ge \mu M_n^+ + \mu(J_n^+ - M_n^+) = \mu(M_n^+ \cap J_n^+) + \mu(J_n^+ - M_n^+) = \mu J_n^+ = h .$$

Let $S_n = \left\{x \varepsilon (x_o - 2h, x_o + 2h) : f(x) > n\right\}$. By (3.7.4) $M_n^+ \subset M_n \subset S_n$. Thus for any positive integer n

$$0 < h < \mu S_n \leq 4h, \tag{3.7.5}$$

since S_n is a measurable set. Hence, since $\left\{S_n\right\}$ is a descending sequence of measurable sets and $\mu S_1 \leq 4h$, $\lim_{n \to \infty} \mu S_n$ exists, and by (3.7.5),

$$\lim_{n \to \infty} \mu S_n \geq h > 0 . \tag{3.7.6}$$

On the other hand, by the continuity of Lebesgue measure,

$$\lim_{n \to \infty} \mu S_n = \mu \lim_{n \to \infty} S_n = \mu \bigcap_{n=1}^{\infty} S_n .$$

By the definition of S_n,

$$\bigcap_{n=1}^{\infty} S_n = \left\{x \varepsilon (x_o - 2h, x_o + 2h) : f(x) = +\infty\right\} = \emptyset ,$$

since f is real-valued. Thus

$$\lim_{n \to \infty} \mu S_n = \mu \emptyset = 0 ,$$

which contradicts (3.7.6).

The assumption that there is a point $x_o \varepsilon (a, b)$ at which f is not continuous leads to a contradiction. Hence f is continuous everywhere on (a, b). This is equivalent to saying that a real-valued discontinuous midpoint convex function on (a, b) cannot be Lebesgue measurable. ∎

<u>Corollary 3.8</u>. Let f be a real-valued function midpoint convex on the finite closed interval $[a, b]$. Then f is continuous on (a, b) if and only if f is measurable on some nonempty open interval $(c, d) \subset (a, b)$.

The corollary is an immediate consequence of Theorem 3.7 and Corollary 3.5.

CHAPTER IV

ANALYSIS OF SOME FUNCTIONAL EQUATIONS

This chapter is primarily concerned with the analysis of the two related functional equations

$$f(x + y) = f(x) + f(y) \tag{4.0.1}$$

and,

$$f(x + y) = f(x) \cdot f(y) . \tag{4.0.2}$$

The study of these equations was initiated by Cauchy (cf. Cauchy [3]). In equation (4.0.1) f is assumed to be a real-valued function defined on the real line, and in equation (4.0.2), f is assumed to be a complex-valued function defined on $[0, \infty)$. It is shown that under the mild restriction of Lebesgue measurability of the function f the general solutions of the functional equations (4.0.1) and (4.0.2) may be obtained. It was proved by Cauchy that if f is assumed to be continuous (or, what is equivalent, continuous at $x = 0$) then (4.0.1) has the general solution $f(x) = cx$ where c is a real constant. It is surprising that the same result applies under the apparently far weaker restriction of Lebesgue measurability. There are several proofs of this latter result in the literature, of varying degrees of difficulty. The original proof is due to M. Fréchet, but the argument used here follows the method of M. Kac. Detailed references are given in the text of this chapter.

Lemma 4.1. If f is a real-valued function with domain E_1 and if f satisfies the functional equation $f(x+y) = f(x) + f(y)$ for all x, y in E_1, then

$$f\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} f(x_i)$$

whenever $x_1, x_2, \ldots, x_n \in E_1$.

The proof, using an obvious induction, is omitted.

Lemma 4.2. If f is a continuous real-valued function satisfying $f(x + y) = f(x) + f(y)$ for all x, y in E_1, then $f(x) = cx$ for some constant c.

Proof. By Lemma 4.1, $f(x) = f(n \frac{x}{n}) = nf(\frac{x}{n})$, $n = 1, 2, \ldots$. If m, n are positive integers,

$$f(\frac{m}{n} x) = mf(\frac{x}{n}) = \frac{m}{n} nf(\frac{x}{n}) = \frac{m}{n} f(x) .$$

Since $f(x) = f(x + 0) = f(x) + f(0)$, $f(0) = 0$. Then $f(x) + f(-x) = 0$, so that $f(-x) = -f(x)$. It follows that, for all rational numbers r, $f(rx) = rf(x)$. Let $f(1) = c$. Then $f(r) = rc$ for all rational numbers r. Since f is continuous, for any real x, and $\{r_n\}$ a sequence of rationals with limit x, it follows that

$$f(x) = \lim_{n \to \infty} f(r_n) = \lim_{n \to \infty} cr_n = cx . \blacksquare$$

In the next theorem, which is due to Hamel [9], it is shown that equation (4.01) can be satisfied by an everywhere discontinuous function.

The proof depends on the notion of a Hamel basis, which briefly is a set of real numbers (a) such that any real number is representable as a linear combination of a finite subset of the basis, with rational coefficients, and (b) such that any rational combination of a finite subset of the basis is zero only if the rational coefficients are all zero. In the APPENDIX Hamel bases are discussed. There it is shown that such bases exist, and a unique representation theorem is proved.

Theorem 4.3. There exists a real-valued function f which is everywhere discontinuous and satisfies the functional equation
$f(x + y) = f(x) + f(y)$ for all x, y in E_1.

Proof. Suppose that H is a Hamel basis for E_1. Let b' and b'' be elements of H, with $b' \neq b''$. Let $f(b')$, $f(b'')$ be real numbers such that

$$\frac{f(b')}{b'} \neq \frac{f(b'')}{b''} \tag{4.3.1}$$

Let $f(b)$ be a real number (arbitrarily assigned) for all other elements $b \in H$; e.g., $f(b) = b$ for all $b \in H - \{b', b''\}$. Let $f(0) = 0$. If $x \in E_1$, $x \neq 0$ then x has exactly one representation of the form

$$x = \sum_{i=1}^{m} r_i b_i \, , \tag{4.3.2}$$

where $b_1, b_2, \ldots, b_m$ are distinct elements of H, $r_1, r_2, \ldots, r_m$ are nonzero rational numbers, and m is a positive integer. Define

$$f(x) = \sum_{i=1}^{m} r_i f(b_i) \quad \text{if} \quad x = \sum_{i=1}^{m} r_i b_i ,$$

using the representation of the form (4.3.2). Then f is a real-valued function defined for all $x \varepsilon E_1$. It will now be shown that $f(x + y) = f(x) + f(y)$ for all $x, y \varepsilon E_1$. This is evident if either $x = 0$ or $y = 0$. Thus suppose $x \neq 0$, $y \neq 0$, and let

$$x = \sum_{i=1}^{m} r_i b_i, \quad y = \sum_{i=1}^{n} s_i c_i ,$$

using representations of the form (4.3.2). Let

$$\{d_1, d_2, \ldots, d_p\} = \{b_1, b_2, \ldots, b_m\} \cup \{c_1, c_2, \ldots, c_n\}.$$

Then

$$x = \sum_{i=1}^{p} r_i' d_i, \quad y = \sum_{i=1}^{p} r_i'' d_i$$

where the numbers r_i', r_i'' are rational (now not in general all nonzero). Then

$$x + y = \sum_{i=1}^{p} (r_i' + r_i'') d_i$$

so that

$$f(x + y) = \sum_{i=1}^{p} (r_i' + r_i'')\; f(d_i) = \sum_{i=1}^{p} r_i' f(d_i) + \sum_{i=1}^{p} r_i'' f(d_i) = f(x) + f(y).$$

On the other hand, if f is continuous at some $x_o \in E_1$, then it follows easily from the identity $f(x + h) = f(x) + f(x_o + h) - f(x_o)$ that f is continuous everywhere in E_1. In the latter case, $f(x) = cx$ for some constant c, as shown previously. But then, using $b', b'' \in H$ as in the first part of the proof, it would follow that $f(b')/b' = f(b'')/b''$. However this contradicts the requirement (4.3.1) imposed in defining f. Hence f is an everywhere discontinuous real-valued function satisfying $f(x + y) = f(x) + f(y)$ for all x, y in E_1. ∎

Remark. The function f constructed in Theorem 4.3 is midpoint convex, since for all x, y in E_1

$$f(x) + f(y) = f(x + y) = f\left(\frac{x+y}{2} + \frac{x+y}{2}\right) = 2f\left(\frac{x+y}{2}\right)$$

and thus

$$f\left(\frac{x+y}{2}\right) = \frac{f(x) + f(y)}{2}.$$

Thus the above theorem provides an example of an everywhere discontinuous real-valued midpoint convex function on $(-\infty, \infty)$.

Remark. The following theorem, due to M. Fréchet [4, 5], is of considerable interest and importance. Several other proofs have appeared in the literature. A proof based on the theory of convex functions can be constructed from Theorems 3.4 and 4.3. The argument used follows the method of M. Kac [15], involving a technique which has further applications (in the solution of some related functional equations).

Theorem 4.4. Let f be a real-valued function which satisfies the equation

$$f(x + y) = f(x) + f(y) \tag{4.4.1}$$

for all real numbers x and y. If f is Lebesgue measurable, then f is continuous on $(-\infty, \infty)$, and $f(x) = cx$ for some real constant c.

Proof. The function $\exp(if(x))$ is a bounded (complex-valued) Lebesgue measurable function, and thus is summable over any set of finite Lebesgue measure. Let a be a real number such that

$$\int_0^a \exp(if(t))\, dt \neq 0 .$$

If there were no such number a then the integrand would be zero almost everywhere, a hypothesis which is false, since $|\exp(if(t))| = 1$ for all real numbers t. Using (4.4.1), it follows that

$$\int_0^a \exp(if(x)) \cdot \exp(if(t))\, dt = \int_0^a \exp(if(x+t))\, dt ,$$

and, using a change of variable theorem for the Lebesgue integral,

$$\exp(if(x)) \int_0^a \exp(if(t))\, dt = \int_x^{x+a} \exp(if(t))\, dt .$$

Hence,

$$\exp(if(x)) = \frac{\int_x^{x+a} \exp(if(t))\, dt}{\int_0^a \exp(if(t))\, dt} . \tag{4.4.2}$$

Since the indefinite integral of a Lebesgue summable function is continuous, it follows from (4.4.2) that $\exp(if(x))$ is continuous for all real x.

It follows easily from the functional equation (4.4.1) that $f(0) = 0$ and

$$f\left(\frac{p}{q}x\right) = \frac{p}{q}f(x) \qquad (4.4.3)$$

for all real numbers x, and all integers p and q $(q \neq 0)$. Let $g(x) = \exp(if(x))$. Then

$$g\left(\frac{p}{q}\right) = \exp\left(i\frac{p}{q}f(1)\right)$$

for all integers p and q $(q \neq 0)$. Given a real number x, let $\{r_n\}$ be a sequence of rationals with limit x. Since g is continuous at x,

$$g(x) = \lim_{n\to\infty} g(r_n) = \lim_{n\to\infty} \exp(ir_n f(1)) = \exp(ixf(1)).$$

Thus, if $c = f(1)$

$$\exp(if(x)) = g(x) = \exp(icx), \ -\infty < x < \infty,$$

or equivalently

$$\exp(i(f(x) - cx)) = 1, \ -\infty < x < \infty.$$

Hence

$$f(x) - cx = 2n(x)\pi$$

where $n(x)$ is an <u>integer</u> for each x. The proof is completed by showing that $n(x) = 0$ for all real x.

Using (4.4.3), and the fact that $f(x) = cx + 2n(x)\pi$,

$$\frac{p}{q}(cx + 2n(x)\pi) = f\left(\frac{p}{q}x\right) = c\left(\frac{p}{q}x\right) + 2n\left(\frac{p}{q}x\right)\pi,$$

and thus

$$n\left(\frac{p}{q}x\right) = \frac{p}{q}n(x) \tag{4.4.4}$$

for all real numbers x and all integers p,q with $q \neq 0$. Since $f(0) = 0$, $n(0) = 0$. Suppose there exists x_o for which $n(x_o) = j \neq 0$, where j is an integer. Then, using (4.4.4)

$$n\left(\frac{p}{q}\frac{x_o}{j}\right) = n\left(\frac{p}{qj}x_o\right) = \frac{p}{qj}j = \frac{p}{q}.$$

This is impossible since $n(x)$ is integer-valued, and the supposition that $n(x_o) \neq 0$ for some x_o leads to a contradiction if $p = 1$, $q = 2$, for example. Hence $n(x) = 0$ for all real x, and thus $f(x) = cx$ for all real x. The continuity and the form of $f(x)$ are deduced at the same time. ∎

The remaining theorems in the chapter are concerned with the functional equation

$$f(x + y) = f(x) \cdot f(y) .$$

The proofs are for the most part detailed versions of those in Hillè and Phillips [12], pp. 144-146.

<u>Theorem 4.5</u>. Let f be a complex-valued function defined on $[0, \infty)$. Let $f(0) = 1$, and $f(x + y) = f(x) \cdot f(y)$ for all $x, y > 0$.

If, for some $x_o > 0$, $f(x_o) \neq 0$ and if $|f(x)|$ is bounded in some interval $[a, b]$, where $0 \leq a < b$, then $|f(x)| = \exp (cx)$ for some real number c.

Proof. Suppose for some $x_1 > 0$ that $f(x_1) = 0$. For $x > x_1$,

$$f(x) = f(x - x_1)\, f(x_1) = 0 .$$

For $0 < x < x_1$, let n be an integer such that $nx > x_1$. Then $(f(x))^n = f(nx) = 0$, since $nx > x_1$. Thus $f(x) = 0$ for all $x > 0$. This is a contradiction, and thus since $f(x_o) \neq 0$, $f(x) \neq 0$ for $x \geq 0$.

Let the function g be defined on $[0, \infty)$ by

$$g(x) = \ln |f(x)| .$$

Notice that $g(x + y) = g(x) + g(y)$, $g(0) = 0$ and that $g(x)$ is bounded above on $[a, b]$. It will be shown that $g(x)$ is actually bounded in a neighborhood of the origin.

Suppose that $g(x)$ is not bounded in any neighborhood of the origin. Then there exists a sequence $\{x_n\}_{n=1}^{\infty}$ such that $0 < x_n < b - a$ for $n \geq 1$, $\lim_{n \to \infty} x_n = 0$, and $\lim_{n \to \infty} |g(x_n)| = +\infty$. Hence

$$\lim_{n \to \infty} |g(a + x_n)| = \lim_{n \to \infty} |g(a) + g(x_n)| = +\infty .$$

Since g is bounded above in $[a, b]$, this implies that

$$\lim_{n \to \infty} g(x_n) = -\infty$$

and hence that

$$\lim_{n \to \infty} \left(g(b) - g(x_n)\right) = \lim_{n \to \infty} \left(g(b - x_n)\right) = +\infty .$$

This contradicts the fact that g is bounded above in $[a, b]$. Thus g is bounded in some neighborhood of the origin.

It will now be shown that

$$\lim_{x \to 0+} g(x) = 0 .$$

If it is false that $\lim_{x \to 0+} g(x) = 0$, then there exists a strictly decreasing sequence $\{y_n\}$ with limit zero such that $\{|g(y_n)|\}$ does not have limit zero. This implies that $\limsup |g(y_n)| = \lambda > 0$. Choose $\varepsilon = \frac{\lambda}{2}$. Then there exists a subsequence $\{z_n\}$ of $\{y_n\}$ such that $|g(z_n)| \geq \varepsilon$ for $n = 1, 2, \ldots$. The sequence $\{z_n\}$ in turn has a subsequence $\{w_n\}$ with the property that either $g(w_n) \geq \varepsilon$ for all $n \geq 1$ or $g(w_n) \leq -\varepsilon$ for all $n \geq 1$. The sequence $\{w_n\}$ is decreasing and has limit zero. Suppose that $g(w_n) \geq \varepsilon$ for all $n \geq 1$. Then, for all positive integers n, k,

$$g\left(\sum_{i=k+1}^{n+k} w_i\right) = \sum_{i=k+1}^{n+k} g(w_i) \geq n\varepsilon .$$

If $g(w_n) \leq -\varepsilon$ for all $n \geq 1$, then similarly

$$g\left(\sum_{i=k+1}^{k+n} w_i\right) \leq -n\varepsilon .$$

Thus

$$\left| g\left(\sum_{i=k+1}^{k+n} w_i\right) \right| \geq n\varepsilon$$

for all positive integers k, n. Since $\{w_i\}$ is a strictly decreasing sequence with limit zero,

$$0 < \sum_{i=k+1}^{k+n} w_i < nw_{k+1} ,$$

and for each fixed n, it follows that

$$\lim_{k \to \infty} \left(\sum_{i=k+1}^{k+n} w_i \right) = 0$$

Thus for each fixed $n = 1,2,\ldots$ there is a sequence $\{p_k\}$ of positive numbers with limit zero such that $|g(p_k)| \geq n\varepsilon$. This contradicts the fact that g is bounded in some neighborhood of the origin. Thus $\lim_{x \to 0+} g(x) = 0$.

Notice that $g(0) = 0$. Thus it has been proved that g is right continuous at the origin. Observe that

$$\lim_{h \to 0+} g(x + h) = \lim_{h \to 0+} (g(x) + g(h)) = g(x) + \lim_{h \to 0+} g(h) = g(x) .$$

Thus g is right continuous at each point $x \geq 0$.

Finally, let r be positive and rational. Then there exist positive integers m and n such that $r = m/n$.

Then

$$g(r) = g(\tfrac{m}{n}) = m\,g(\tfrac{1}{n}) = \tfrac{m}{n}\, n\, g(\tfrac{1}{n}) = \tfrac{m}{n}\, g(1) ,$$

and thus $g(r) = rg(1)$. For any $x \geq 0$ there is a sequence $\{r_n\}_1^\infty$ of rationals such that $r_n \longrightarrow x^+$. Then since g is right continuous at x, and $g(r_n) = r_n g(1)$, it follows that

$$g(x) = g\,(\lim_{n \to \infty} r_n) = \lim_{n \to \infty} g(r_n) = \lim_{n \to \infty} r_n g(1) = xg(1) .$$

Thus $|f(x)| = \exp(g(1)x)$ for $x \geq 0$. ∎

Definition 4.6. Let f be a complex-valued function defined on $[0, +\infty)$ such that

$$f(x + y) = f(x) \cdot f(y)$$

for all $x, y \geq 0$,

$$f(0) = 1$$

and, for all $x \geq 0$,

$$f(x) \neq 0 .$$

Define the function K by

$$K(x) = \frac{f(x)}{|f(x)|} \quad \text{if} \quad x \geq 0 ,$$

$$= \frac{1}{K(-x)} \quad \text{if} \quad x < 0 .$$

Then K is a character of the real line.

Remark. Notice that $K(x + y) = K(x) \cdot K(y)$ for all real x and y, and $|K(x)| = 1$.

Lemma 4.7. Let f be a real-valued function Lebesgue summable on $[c, d]$ and $-\infty \leq c < a < b < d \leq \infty$. Then,

$$\lim_{h \to 0} \int_a^b |f(x + h) - f(x)| \, dx = 0 .$$

Proof. Part (i). Assume, in this part of the proof, that f is continuous on E_1, and, for some integer N, $f(x) = 0$ for each

$x \varepsilon E_1 - [-N, N]$. Then $|f(x)|$ has a maximum value K on $[-N, N]$. Let $\{h_n\}_1^\infty$ be a sequence such that $|h_n| \le 1$ for $n \ge 1$ and

$$\lim_{n\to\infty} h_n = 0.$$

Then $|f(x + h_n) - f(x)| \le 2K$ for all $x \varepsilon [-N - 1, N + 1]$. Now in this case

$$\int_{-\infty}^{\infty} |f(x + h_n) - f(x)|dx = \int_{-N-1}^{N+1} |f(x+h_n) - f(x)| \, dx .$$

By the Lebesgue dominated convergence theorem,

$$\lim_{n\to\infty} \int_{-N-1}^{N+1} |f(x + h_n) - f(x)|dx = \int_{-N-1}^{N+1} \lim_{n\to\infty} |f(x+h_n) - f(x)| \, dx.$$

But

$$\lim_{n\to\infty} |f(x + h_n) - f(x)| = 0, \ -\infty < x < \infty,$$

since f is continuous on E_1.

Hence

$$\lim_{h\to 0} \int_{-\infty}^{\infty} |f(x+h) - f(x)| \, dx = 0 .$$

Part (ii). Define the function $\tilde{f}$ by

$$\begin{aligned} \tilde{f}(x) &= f(x), \ x \varepsilon [c, d] \\ &= 0 \ , \ x \varepsilon E_1 - [c, d]. \end{aligned}$$

Then, since f is summable on $[c, d]$, $\tilde{f}$ is summable on E_1, and

$$\int_{-\infty}^{\infty} \tilde{f}(x)\, dx = \int_{c}^{d} f(x)\, dx$$

By a standard approximation theorem for summable functions (cf. Hewitt [11], pp. 143-4, or Groome [7], pp. 11-15, for detailed proofs), given $\varepsilon > 0$ there exists a function g continuous on E_1 such that

$$\int_{-\infty}^{\infty} |\tilde{f}(x) - g(x)|\, dx < \frac{\varepsilon}{3}$$

and a positive integer N such that

$$g(x) = 0, \quad x \in E_1 - [-N, N].$$

Note that

$$|\tilde{f}(x + h) - \tilde{f}(x)| = |\tilde{f}(x + h) - \tilde{f}(x) - g(x+h) + g(x) + g(x+h) - g(x)|$$
$$\leq |\tilde{f}(x + h) - g(x+h)| + |\tilde{f}(x) - g(x)| + |g(x+h) - g(x)|$$

for all real x and h.

By Part (i), given $\varepsilon > 0$ there exists an $h' > 0$ such that for $|h| < h'$

$$\int_{-\infty}^{\infty} |g(x + h) - g(x)|\, dx < \frac{\varepsilon}{3}.$$

Now

$$\int_{-\infty}^{\infty} |\tilde{f}(x + h) - g(x + h)|\, dx = \int_{-\infty}^{\infty} |\tilde{f}(x) - g(x)|\, dx < \frac{\varepsilon}{3}$$

using a change of variable theorem for the Lebesgue integral. Thus, for $|h| < h'$,

$$\int_{-\infty}^{\infty} |\tilde{f}(x + h) - \tilde{f}(x)|dx \leq \int_{-\infty}^{\infty} |\tilde{f}(x + h) - g(x + h)|dx$$

$$+ \int_{-\infty}^{\infty} |\tilde{f}(x) - g(x)|dx +$$

$$+ \int_{-\infty}^{\infty} |g(x+h) - g(x)|dx < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Hence

$$\lim_{h \to 0} \int_{-\infty}^{\infty} |\tilde{f}(x + h) - \tilde{f}(x)|\, dx = 0.$$

Since, $\tilde{f}(x) = f(x)$ for $x \in [c, d]$ and since there is some $h'' > 0$ such that $x + h \in [c, d]$ if $x \in [a, b]$ and $|h| < h''$,

$$\int_{-\infty}^{\infty} |\tilde{f}(x + h) - \tilde{f}(x)|\, dx \geq \int_{a}^{b} |f(x + h) - f(x)|dx \geq 0.$$

Thus

$$\lim_{h \to 0} \int_{a}^{b} |f(x + h) - f(x)|dx = 0. \blacksquare$$

<u>Theorem 4.8</u>. If K is a measurable character of the real line, then $K(x) = \exp(i\, b\, x)$ for some real number b.

<u>Proof</u>. For $|h| > 0$ and x and y real,

$$K(x + h) - K(x) = K(y)(K(x+h-y) - K(x-y)).$$

Thus for $r > 0$

$$\int_{0}^{r} (K(x+h) - K(x))dy = \int_{0}^{r} K(y)(K(x+h-y) - K(x-y))dy$$

and

$$K(x + h) - K(x) = \frac{1}{r}\int_0^r K(y)(K(x + h - y) - K(x - y))dy .$$

Hence, since $|K(y)| = 1$ for all real y,

$$|K(x + h) - K(x)| \leq \frac{1}{r}\int_0^r |K(x+h-y) - K(x - y)|\, dy . \quad (4.8.1)$$

Thus

$$\lim_{h \to 0} |K(x + h) - K(x)| = 0 ,$$

since the limit on the right in inequality (4.8.1) is zero by Lemma 4.7. Thus K is continuous for all real x.

Considering $r > 0$,

$$\frac{K(h) - K(0)}{h} \int_0^r K(x)\, dx = \frac{1}{h} \int_0^r K(x + h) - K(x)\, dx ,$$

$$= \frac{1}{h} \int_0^r K(x + h)dx - \frac{1}{h}\int_0^r K(x)dx$$

$$= \frac{1}{h}\int_h^{r+h} K(x)\, dx - \frac{1}{h}\int_0^r K(x)\, dx .$$

Denoting H such that $H'(x) = K(x)$ on E_1, it follows by the fundamental theorem of calculus that

$$\int_h^{r+h} K(x)dx - \int_0^r K(x)\, dx = H(r + h) - H(h) - H(r) + H(0),$$

$$= \int_r^{r+h} K(x)\, dx - \int_0^h K(x)\, dx .$$

Thus

$$\frac{K(h) - K(0)}{h} \int_0^r K(x)\,dx = \frac{1}{h} \int_r^{r+h} K(x)\,dx - \frac{1}{h} \int_0^h K(x)\,dx .$$

Since K is continuous,

$$\lim_{h \to 0} \frac{1}{h} \int_r^{r+h} K(x)\,dx = K(r) ,$$

and

$$\lim_{h \to 0} \frac{1}{h} \int_0^h K(x)\,dx = K(0) .$$

Since $\int_0^r K(x)\,dx \neq 0$ for all $r > 0$, there exists $r > 0$ such that

$$\lim_{h \to 0} \frac{K(h) - K(0)}{h} = \frac{K(r) - K(0)}{\int_0^r K(x)\,dx}$$

and K is differentiable at the origin.

Let

$$b = -iK'(0) .$$

Then

$$\frac{K(x + h) - K(x)}{h} = K(x) \left(\frac{K(h) - K(0)}{h} \right)$$

and thus $K'(x) = i\,b\,K(x)$, for each real number x. Therefore, for some constant C, $K(x) = C \exp(i\,b\,x)$ for all real x. Using the fact that $K(0) = 1$, it follows that $C = 1$ and $K(x) = \exp(i\,b\,x)$. If $b = b_1 + ib_2$, (where b_1 and b_2 are real), then

$$\exp(ibx) = \exp(ib_1 x - b_2 x)$$

and

$$|\exp(ibx)| = \exp(-b_2 x) .$$

If $b_2 \neq 0$ then $|\exp(ibx)|$ would be unbounded on E_1 contradicting the fact that $|K(x)| = 1$. Thus there is some real number b such that

$$K(x) = \exp(ibx) . \blacksquare$$

Corollary 4.9. Let f be a complex-valued Lebesgue measurable function defined on $[0, \infty)$ such that $f(0) = 1$, $f(x + y) = f(x) \cdot f(y)$ for each nonnegative x and y, and $f(x) \neq 0$ for all $x \geq 0$. Then there are real numbers a and b such that

$$f(x) = \exp((a + ib) x)$$

for all $x \geq 0$.

Proof. Let g be defined by

$$g(x) = \ln |f(x)| .$$

Then g is finite-valued, measurable and

$$g(x + y) = g(x) + g(y)$$

for all real $x, y \geq 0$. Since, for all real $x, y \geq 0$,

$$g(x) + g(y) = g(x + y) = g\left(\tfrac{1}{2} (x+y) + \tfrac{1}{2} (x+y)\right) = 2g\left(\tfrac{1}{2} (x+y)\right) ,$$

it follows that g is midpoint convex on $(0, \infty)$. By Theorem 3.7, g

is continuous on $(0, \infty)$ and, thus, bounded on each interval of the form $[\varepsilon, \varepsilon^{-1}]$ for $0 < \varepsilon < 1$. By Theorem 4.5, $|f(x)| = \exp(ax)$ for some real number a. If K is defined by

$$K(x) = \frac{f(x)}{|f(x)|},$$

then K is measurable. Thus

$$\frac{f(x)}{|f(x)|} = \frac{f(x)}{\exp(ax)} = K(x) = \exp(ibx)$$

for some real number b. Therefore

$$f(x) = \exp(ax)\,\exp(ibx) = \exp((a + ib)x)\,. \blacksquare$$

APPENDIX

HAMEL BASES

Definition A.1. A Hamel basis for the set of real numbers (viewed as a vector space over the rational field) is a set H of real numbers which satisfies the following two conditions.

(i). Given a real number x, x can be represented in the form

$$x = \sum_{i=1}^{n} r_i h_i$$

where n is a positive integer, $\{h_1, h_2, \ldots, h_n\} \subset H$, and $r_1, r_2, \ldots, r_n$ are rational numbers.

(ii). Whenever $h_1, h_2, \ldots, h_n$ (for some positive integer n) are elements of H with $h_i \neq h_j$ for $i \neq j$, and $r_1, r_2, \ldots, r_n$ are rational numbers,

$$\sum_{i=1}^{n} r_i h_i = 0$$

only if $r_i = 0$, $i = 1, 2, \ldots, n$.

Condition (ii) asserts that H is a finitely linearly independent set in the vector space concerned. The theorem which follows establishes the existence of a Hamel basis for the reals. The result was first proved by Hamel [9] in 1905 using the well-ordering theorem. The proof given here makes use of Zorn's lemma which is equivalent to the well-ordering theorem. For the terminology used, consult Kelley [16] or McShane and Botts [21].

Theorem A.2. There exists a Hamel basis for the set E_1 of real numbers.

Proof. Let $\mathcal{P}$ be the collection of all subsets of E_1 which are finitely linearly independent in the sense of condition (ii) of Definition A.1. The collection $\mathcal{P}$ is nonempty, since, for example, the one-point set $\{2\}$ belongs to $\mathcal{P}$. The collection $\mathcal{P}$ is partially ordered by set inclusion. Suppose that $\mathcal{C}$ is a chain in $\mathcal{P}$, i.e., a nonempty subcollection of $\mathcal{P}$ such that whenever $P \varepsilon \mathcal{C}$ and $Q \varepsilon \mathcal{C}$ then either $P \subset Q$ or $Q \subset P$. Let C be the union of the sets in the chain $\mathcal{C}$. It is shown in what follows that C is itself a finitely linearly independent set, and consequently belongs to $\mathcal{P}$ and is an upper bound in $\mathcal{P}$ for the chain $\mathcal{C}$.

Suppose that C is not a finitely linearly independent set. Then C contains a finite subset $\{c_1, c_2, \ldots, c_n\}$ which is linearly dependent. Let $c_i \varepsilon C_i \varepsilon \mathcal{C}$, $i = 1,2,\ldots,n$. Since $\mathcal{C}$ is a chain, for each i and j either $C_i \subset C_j$ or $C_j \subset C_i$. Consequently there exists an index $m \varepsilon \{1,2,\ldots,n\}$ for which $C_i \subset C_m$, $i = 1,2,\ldots,n$. It follows that $\{c_1, c_2, \ldots, c_n\} \subset C_m$. This contradicts the fact that C_m, as a member of $\mathcal{P}$, is a finitely linearly independent set. Hence C is a member of $\mathcal{P}$, and C clearly is an upper bound for $\mathcal{C}$.

Every chain in $\mathcal{P}$ has an upper bound in $\mathcal{P}$. Zorn's lemma implies that $\mathcal{P}$ contains a maximal element H. This asserts that H is a finitely linearly independent subset of E_1 which is not properly contained in any other finitely linearly independent subset of E_1. Such a maximal set H is a Hamel basis for E_1, as is seen by the following. It is only necessary to verify condition (i) of the definition. Suppose that $x \in E_1$ but $x \notin H$. By the maximality of H, the set $H \cup \{x\}$ is a finitely linearly dependent set. Otherwise the maximality of H is contradicted, since $H \cup \{x\}$ properly contains H. Hence there exist rational numbers $a_1, a_2, \ldots, a_k$, not all zero, for which

$$a_1 h_1 + \ldots + a_{k-1} h_{k-1} + a_k x = 0 .$$

Here, $a_k \neq 0$ since if $a_k = 0$, then H would be finitely linearly dependent, which is not so. Thus $a_k \neq 0$, and x can be represented in the form required by condition (i) of the definition. It follows that H is a Hamel basis for E_1. ∎

Note. It is easy to verify that a Hamel basis for E_1 is non-denumerable, since the collection of all finite linear combinations (with rational coefficients) of a denumerable set is denumerable. A Hamel basis for E_1 can contain at most one rational number, and there exists a Hamel basis which contains the rational number 1. If a Hamel basis H were to contain two rational numbers, then there would be a finitely linearly dependent subset of H, which is not so for any Hamel basis. If H is a Hamel basis for the set of real numbers, it is easily verified that the set $H_1 = \left\{ \frac{h}{h_1} : h \in H \right\}$, formed for a fixed element $h_1 \in H$, is a Hamel basis which contains the rational number 1.

Theorem A.3. Given a Hamel basis H for E_1 every nonzero $x \in E_1$ has a unique representation of the form

$$x = \sum_{i=1}^{n} r_i h_i ,$$

where $h_1, h_2, \ldots, h_n$ are distinct elements of H and the rational coefficients $\{r_i\}$ are nonzero.

Proof. Suppose that $x \varepsilon E_1$, $x \neq 0$ and that

$$x = \sum_{i=1}^{n} r_i a_i$$

and

$$x = \sum_{j=1}^{p} s_j b_j \quad (1 \leq n \leq p) ,$$

where $a_1, a_2, \ldots, a_n$ are distinct elements of H; $b_1, b_2, \ldots, b_p$ are distinct elements of H; and, the rational numbers $r_1, r_2, \ldots, r_n$; $s_1, s_2, \ldots, s_p$ are nonzero. It must be shown that $n = p$, $\{a_1, a_2, \ldots, a_n\} = \{b_1, b_2, \ldots, b_p\}$, and that $r_i = s_i$, $i = 1, 2, \ldots, n$.

Suppose that for some $m \geq 1$, $\{b_1, b_2, \ldots, b_m\} \subset \{a_1, a_2, \ldots, a_n\}$. It may be assumed that the elements in each set are so labeled that

$$a_1 = b_1, \quad a_2 = b_2, \ldots, \; a_m = b_m.$$

If $m < n$ and $m < p$, then

$$x = \sum_{i=1}^{m} r_i a_i + \sum_{i=m+1}^{n} r_i a_i ,$$

and

$$x = \sum_{i=1}^{m} s_i a_i + \sum_{i=m+1}^{p} s_i b_i .$$

It follows that

$$0 = \sum_{i=1}^{m} (r_i - s_i)\, a_i + \sum_{i=m+1}^{n} r_i a_i + \sum_{i=m+1}^{p} (-s_i) b_i ,$$

where all the a_i, b_i are distinct elements of H, and the rational numbers $r_1, r_2, \ldots, r_n$; $s_1, s_2, \ldots, s_p$ are nonzero. This is impossible since H is a finitely linearly independent set. Similarly, it is impossible to have $m < n$, $m = p$ or $m = n$, $m < p$. Thus necessarily $m = n$ and $m = p$. In this case,

$$0 = \sum_{i=1}^{n} (r_i - s_i) a_i$$

and thus $r_i = s_i$, $i = 1, 2, \ldots, n$. ∎

Sierpiński [26] has given an ingenious modification of the proof of existence of a Hamel basis for the real numbers wherein he shows that there exists a Hamel basis of zero Lebesgue measure. The idea of Sierpiński's argument will be indicated only, without details. Let X be the set of all real numbers x which admit binary expansions of the form

$$x = C_0 \cdot C_1 0 C_3 0 C_5 0 \ldots$$

where C_0 is an integer and $C_{2i-1} = 0$ or 1 $(i = 1, 2, \ldots.)$. Let Y be the set of all real numbers x which admit binary expansions of the form

$$x = 0 \cdot 0C_2 0C_4 0C_6 \cdots$$

where $C_{2i} = 0$ or 1 $(i = 1,2,\ldots.)$. The sets X and Y are sets of zero Lebesgue measure, as is the set $S = X \cup Y$. Every real number z can be represented uniquely in the form $z = x + y$ where $x \in X$ and $y \in Y$. The argument used in the usual existence proof for a basis is then applied to subsets of S which are finitely linearly independent. In view of the representation $z = x + y$, a Hamel basis for E_1 is obtained. This basis is a subset of S and hence has Lebesgue measure zero. For a more detailed discussion (using well-ordering in the proof, instead of Zorn's lemma), consult Sierpiński [26].

Sierpiński further proves that, if a Hamel basis H for the real numbers is Lebesgue measurable, then H must be of measure zero. Suppose $\mu H > 0$, and let $h_1 \in H$. Consider the set

$$H_1 = \left\{ \frac{h}{h_1} : h \in H \right\}.$$

Then

$$\mu H_1 = \frac{1}{h_1} \mu H > 0$$

by Lemma 3.6. Then H_1 contains distinct elements x and y such that $x - y$ is rational. To deduce this, suppose that H_1 does not contain two distinct elements whose difference is rational. Since $\mu H_1 > 0$ there exists an interval $(-k, k)$ such that $\mu H_1 \cap (-k,k) > 0$. Let $H_2 = H_1 \cap (-k,k)$. Then the sets

$$\left\{h : h = h' + \frac{1}{n},\ h' \varepsilon H_2\right\} = H_2 + \frac{1}{n},\ n = 1, 2, \ldots,$$

are pairwise disjoint, measurable sets, each of measure μH_2, and all sets

$$H_2 + \frac{1}{n}$$

are subsets of $(-k, k + 1)$. Then on one hand

$$\mu \bigcup_{n=1}^{\infty} (H_2 + \frac{1}{n}) \leq 2k + 1 .$$

But

$$\mu \bigcup_{n=1}^{\infty} (H_2 + \frac{1}{n}) = \sum_{i=1}^{\infty} \mu(H_2 + \frac{1}{n}) = +\infty ,$$

and a contradiction is obtained.

Thus H_1 contains distinct elements x and y such that $x - y$ is rational. Let x and y, be two such distinct elements of H_1. Thus

$$x = \frac{h'}{h_1} , \quad y = \frac{h''}{h_1} , \quad h'' \neq h' .$$

Then there is a rational number r such that

$$x - y = \frac{h' - h''}{h_1} = r \neq 0 .$$

Hence

$$1h' - 1h'' - rh_1 = 0 .$$

This contradicts the fact that H is a Hamel basis. Hence if H is

Lebesgue measurable, then $\mu H = 0$.

Although we shall not pursue the matter here, it is of interest to note that there exists a Hamel basis which is not Lebesgue measurable. Furthermore, no Hamel basis is a Borel set. Proofs of these profound assertions are due to W. Sierpiński [26].

BIBLIOGRAPHY

1. Bernstein, F. and Doetsch, G., "Zur Theorie der konvexen Funktionen," Mathematische Annalen, 76, 1915, pp. 514 - 526.

2. Blumberg, H., "On convex functions," Transactions of the American Mathematical Society, 20, 1919, pp. 40 - 44.

3. Cauchy, A. L., Cours d'Analyse de l'École Royal Polytechnique, Ire partie, Analyse algébrique, Paris, 1821 (Oeuvres complétes, IIe serie, III).

4. Fréchet, M., "A propos d'un article sur l'équation fonctionnelle $f(x + y) = f(x) + f(y)$," L'Enseignement Mathématique, 16, 1914, p. 136.

5. Fréchet, M., "Pri la funkcia ekvacio $f(x + y) = f(x) + f(y)$," L'Enseignement Mathématique, 15, 1913, pp. 390 - 393.

6. Galvani, L., "Sulle funzioni convesse di una o due variabila definite in un aggregato qualunque." Rendiconti del Circolo Matematico di Palermo, 41, 1916, pp. 103 - 134.

7. Groome, G. M., Jr., Approximation and Representation Theorems in L_p Spaces, M. S. Thesis, Georgia Institute of Technology, 1964.

8. Hahn, H. and Rosenthal, A., Set Functions. Albuquerque, The University of New Mexico Press, 1948.

9. Hamel, G., "Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x + y) = f(x) + f(y)$," Mathematische Annalen, 60, 1905, pp. 459 - 462.

10. Hardy, G. H., Littlewood, J. E. and Pólya, G., Inequalities, Cambridge, Cambridge University Press, Second Edition, 1952.

11. Hewitt, E., Theory of Functions of a Real Variable, New York, Holt, Rinehart and Winston, 1960.

12. Hille, B., and Phillips, R. S., Functional Analysis and Semi-Groups, Revised Edition, Providence, American Mathematical Society, 1957.

13. Hölder, O., "Über einen Mittelwertsatz," Gesellschaft der Wissenschaften zu Göttingen, 1889, pp. 38 - 47.

14. Jensen, J. L. W. V., "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," Acta Mathematica, 30, 1906, pp. 175-193.

15. Kac, M., "Une remarque sur les équations fonctionnelles," Commentarii Mathematici Helvetici, 9, 1936 - 1937, pp. 170 - 171.

16. Kelley, J. L., General Topology, Van Nostrand, New York, 1955.

17. Krosnosel.skiǐ, M. A., and Ruleckiǐ, Ya. B., Convex Functions and Orlicz Spaces, Groningen, P. Noordhoff Ltd., 1961.

18. Kullback, S., Information Theory and Statistics, New York, Wiley, 1959.

19. Loève, M., Probability Theory, Second Edition, Princeton, D. Van Nostrand, 1960.

20. Marcus, S., "Fonctions Convexes et Fonctions Internes," Bulletin des Sciences Mathématiques, 81, 1957, pp. 66 - 70.

21. McShane, E. J., and Botts, T. A., Real Analysis, Princeton, Van Nostrand, 1959.

22. Montel, P., "Sur les fonctions convexes et les fonctions sousharmoniques," Journal de Mathématiques Pures et Appliquées, Series 9, 7, pp. 29-60.

23. Natanson, I. P., Theory of Functions of a Real Variable, 2 vols., New York, Frederick Ungar, Vol. 1, 1955, Vol. 2, 1961.

24. Sierpiński, W., "Sur l'équation fonctionnelle $f(x + y) = f(x) + f(y)$," Fundamenta Mathematicae, 1, 1920, pp. 116-122.

25. Sierpiński, W., "Sur les fonctions convexes mesurables," Fundamenta Mathematicae, 1, 1920, pp. 125-129.

26. Sierpiński, W., "Sur la question de la base de M. Hamel," Fundamenta Mathematicae, 1, 1920, pp. 105-111.